

UNIVERSIDADE DE LISBOA  
FACULDADE DE CIÊNCIAS  
DEPARTAMENTO BIOLOGIA VEGETAL



**Caracterização Genética da População Imigrante Brasileira  
Residente em Lisboa com Marcadores Genéticos do Tipo InDel  
e do Sistema ESS**

Fátima Carolina da Silva Reis

**Mestrado em Biologia Molecular e Genética**

Dissertação orientada por:  
Prof. Doutor Manuel do Carmo Gomes  
Mestre Cláudia Vieira da Silva

“Talvez não tenha conseguido fazer o melhor, mas lutei para que o  
melhor fosse feito. Não sou o que deveria ser, mas Graças a Deus,  
não sou o que era antes.”

Marthin Luther King

## **Agradecimentos**

Concluída esta etapa da minha caminhada, não posso deixar de agradecer a todos aqueles que a enriqueceram, aos que me acompanharam e motivaram, aos que não me deixaram baixar os braços e me inspiraram até ao último ponto.

A Deus, por estar acima de tudo e de todos. Ao que conhece o passado, o presente e o futuro.

À Faculdade de Ciências, da Universidade de Lisboa, particularmente ao Prof. Doutor Manuel do Carmo Gomes, por ter aceite a orientação da minha dissertação, pela sua preciosa ajuda e pela disponibilidade na resolução das questões que foram surgindo.

À minha orientadora, Mestre Cláudia Vieira da Silva, por ter aceite a orientação da presente tese de Mestrado, pelos incentivos dispensados ao longo da realização da mesma, pela cordialidade, disponibilidade e por todos os conhecimentos transmitidos.

Ao Professor António Amorim, pela oportunidade de ter estagiado no Serviço de Genética e Biologia Forenses da Delegação Sul do Instituto Nacional de Medicina Legal e Ciências Forenses, por ter acompanhado o meu trabalho ao longo da minha passagem pelo Serviço, pela sua amizade e por todos os incentivos e conhecimentos transmitidos.

À Dra. Maria João Porto por ter permitido a realização do presente trabalho no Serviço de Biologia e Genética Forenses.

À Dra. Teresa Ribeiro por me ter acolhido no Serviço de Biologia e Genética Forenses da Delegação Sul do Instituto Nacional de Medicina Legal e Ciências Forenses, bem como pela simpatia e conhecimentos transmitidos.

À Dra. Heloísa Afonso Costa pela sua amizade e auxílio na realização do presente trabalho, bem como pela transmissão de conhecimentos.

A TODA a equipa do Serviço de Biologia e Genética Forenses pela ajuda técnica prestada, por todos os conhecimentos partilhados e pelos bons momentos passados.

À minha família que me tem acompanhado ao longo do meu percurso de vida, pela força e todas as palavras de motivação, por acreditarem em mim e por estarem comigo mesmo quando não é fácil.

À minha mãe, a quem devo a vida e o que hoje sou, por me ter motivado a concluir esta etapa quando tudo parecia não ter fim, pela paciência e por sempre lembrar-me como superar quaisquer batalhas.

A todos, mais uma vez apresento, os meus mais sinceros agradecimentos.

## Resumo

Fortemente relacionados com a globalização, os eventos migratórios estão profundamente enraizados nas sociedades contemporâneas, em virtude da revolução nos transportes e comunicações. É bem conhecido que a migração é um dos principais fatores de variabilidade genética dentro das populações. Nas últimas décadas, Portugal, até então tipicamente um país de emigração, abraçou uma nova realidade, acolhendo imigrantes do Brasil, África, Europa Central e de Leste. Os imigrantes brasileiros são a principal comunidade estrangeira em Portugal, com cerca de 85 000 indivíduos em 2017.

A população brasileira altamente miscigenada - principalmente por Europeus, Africanos e Nativos Americanos - é um reflexo da sua história cultural, começando com a colonização Portuguesa do Brasil, quando cerca de 500 000 colonizadores contactaram com ameríndios locais e cerca de 4 milhões de escravos da África subsariana.

Atualmente, o método generalizado para análise de DNA em contexto forense usa STRs – *Short Tandem Repeats*. O perfil de DNA baseado em STRs permite a definição de perfis genéticos individuais para identificação humana, sendo ao mesmo tempo útil para estudos populacionais, onde certas características genéticas estão associadas a populações ou etnias específicas, ou mesmo a nativos de determinadas regiões. Desde 2009, cinco novos *loci* foram adicionados ao painel Europeu de marcadores genéticos, ESS – *European Standard Set* –, estabelecido para facilitar o intercâmbio de perfis de bancos de dados de DNA entre laboratórios.

Os polimorfismos de inserção/deleção - InDels – caracterizam-se pela presença ou ausência de pequenas sequências de DNA. Estas variações constituem um grupo de marcadores genéticos com vantagens para identificação forense, especialmente em amostras biológicas altamente degradadas, devido ao pequeno tamanho dos fragmentos de amplificação. Além disso, os InDels não só são úteis para resolver casos complexos de parentesco biológico, mas também têm taxas de mutação mais baixas. Portanto, não é de estranhar que os InDels tenham sido apresentados como marcadores moleculares alternativos para aplicação forense, complementarmente ao STRs.

Atualmente, não existem dados disponíveis para marcadores InDel, nem para os últimos *loci* STR do sistema ESS, com foco na população imigrante Brasileira residente em Lisboa. Assim, o principal objetivo deste estudo é caracterizar geneticamente a referida população, com marcadores do tipo InDel e STRs do último painel Europeu de marcadores genéticos.

Os resultados obtidos, demonstraram que todos os *loci* STR estudados, são altamente polimórficos e, excetuando os *loci* TH01 e D21S11, todos se encontram em equilíbrio de Hardy-Weinberg. Um alelo microvariante raro 14.1 foi detetado no *locus* D2S411, com uma frequência de 0,002. Adicionalmente, os resultados para os marcadores InDel, confirmaram a existência de diferenças genéticas entre a população caucasiana Portuguesa, populações imigrantes do Brasil e PALOPs.

**Palavras-chave:** InDels; STRs; ESS; Genética Populacional; Imigrantes brasileiros

## Abstract

Strongly related to globalization, migratory events are deeply rooted in contemporary societies, by virtue of the ongoing revolution in transportation and communication. It is well known that these migratory events are responsible for genetic variability within populations. In recent decades, Portugal, until recently a typically emigration country, embraced a new verity, welcoming immigrants from Brazil, Africa, Central and Eastern Europe. Brazilian immigrants are the main foreign community in Portugal, with about 85 000 individuals in 2017.

The highly admixed Brazilian population - mostly by European, African and Native American populations -, is a reflection of its' cultural history, starting with the Portuguese colonization of Brazil, when around 500 000 colonizers met the local Amerindians and about 4 million Sub-Saharan African slaves.

Presently, the generalized method for DNA analysis in forensic context uses Short Tandem Repeats – STRs. STR-based DNA-profiling enables the definition of individual genetic profiles for human identification, while at the same time being useful for population studies, where certain genetic traits are found associated with specific populations or ethnicities, or even natives of specific regions. Since 2009, five new loci have been added to the European Standard Set – ESS – of Short Tandem Repeats (STRs), established to facilitate the exchange of DNA database profiles between laboratories.

Insertion/deletion polymorphisms – InDels – are characterized by the presence or absence of small DNA sequences. These variations constitute a group of genetic markers with advantages for forensic identification, especially in highly degraded biological samples, where sometimes one needs to analyse small amplification fragments. Furthermore, InDels not only are useful to solve complex cases of biological kinship, but they also have lower mutation rates. Therefore, it is no wonder that InDels have been presented as alternative molecular markers for forensic application, combined with STRs.

Currently, there is no available data for InDel markers, nor for the latest STR *loci* on the ESS panel, with focus on the Brazilian immigrant population living in Lisboa. Thus, the main objective of this study is to genetically characterize the aforementioned population, with InDel markers and STRs from the latest European panel of genetic markers.

The obtained results showed that all STR loci are highly polymorphic and, apart from TH01 and D21SA11, all meet Hardy-Weinberg equilibrium expectations. A rare 14.1 microvariant allele was detected at locus D2S411, with a frequency of 0,002. What is more, the results obtained for InDel markers confirm genetic differences between Portuguese Caucasian population, Brazilian and PALOP's immigrant populations.

**Keywords:** InDels; STRs; ESS; Population genetics; Brazilian immigrants

## **Comunicações Científicas baseadas nesta dissertação**

### **Posters**

Fátima Reis, Cláudia Vieira da Silva, António Amorim, Vanessa Bogas, Teresa Ribeiro, Maria João Porto, Heloísa Afonso Costa (2017) **Genetic characterization of the Brazilian immigrant population in Lisboa with InDel genetic markers**. ISFG 2017. Seul.

### **Artigo Científico**

Fátima Reis, Cláudia Vieira da Silva, António Amorim, Vanessa Bogas, Teresa Ribeiro, Maria João Porto, Heloísa Afonso Costa (2017), **“Genetic characterization of the Brazilian immigrant population in Lisboa with InDel genetic markers”**, *Forensic Science International: Genetics Supplement Series*, 2017. DOI: 10.1016/j.fsigss.2017.09.123

# Índice

Agradecimentos.....	III
Resumo.....	IV
Abstract .....	V
Comunicações Científicas baseadas nesta dissertação .....	VI
Índice de figuras .....	IX
Índice de tabelas .....	XI
Capítulo 1 - Introdução.....	1
1.1    A geografia dos genes .....	1
1.2    Globalização e os fenómenos migratórios.....	1
1.3    Imigração Brasileira em Portugal.....	2
1.3.1 República Federativa do Brasil .....	3
1.4    Genética Populacional .....	4
1.4.1 Equilíbrio de Hardy-Weinberg .....	5
1.4.2 <i>Linkage Dis(equilibrium)</i> .....	5
1.5    Genética Forense .....	6
1.6    Marcadores moleculares em genética forense.....	7
1.6.1 Marcadores do tipo InDel.....	7
1.6.1.1 – Sistema Investigator® DIPlex .....	8
1.6.2 Marcadores do tipo STR.....	8
1.6.3 <i>European Standard Set (ESS)</i> .....	9
1.7    Objetivos .....	10
Capítulo 2 – Materiais e Métodos .....	11
2.1    Procedimento experimental.....	11
2.2    Colheita e armazenamento das amostras .....	11
2.3    Amostragem .....	11
2.4    Extração de DNA .....	12
2.5    Quantificação de DNA .....	12
2.6    Amplificação por PCR <i>multiplex</i> .....	13
2.6.1 Estudo de marcadores do tipo InDel .....	14

2.6.2 Estudo de marcadores do tipo STR .....	14
2.7 Eletroforese Capilar do produto amplificado .....	15
2.8 Detecção e análise do produto amplificado.....	16
2.9 Análise estatística.....	17
2.10 Análise Filogenética .....	17
Capítulo 3 – Resultados e Discussão.....	17
3.1 Estudo dos Marcadores InDel .....	17
3.1.1 Microvariantes.....	19
3.1.1.1 – Locus HLD97 .....	19
3.1.1.2 – Loci HLD99 e HLD56.....	20
3.2 Estudo dos marcadores STR.....	21
3.2.1 Microvariantes <i>off-ladder</i> .....	21
3.3 Estudo da população imigrante Brasileira residente em Lisboa.....	22
3.4 Análise de distâncias genéticas entre populações .....	26
3.5 Análise de <i>Linkage Disequilibrium</i> .....	28
3.6 Análise dos parâmetros Forenses .....	28
Capítulo 4 – Conclusão .....	30
Referências Bibliográficas .....	31
Capítulo 5 – Anexos.....	39
5.1 Características dos <i>loci</i> incluídos nos Sistemas <i>Investigator</i> <sup>®</sup> <i>DIPplex</i> e <i>PowerPlex</i> <sup>®</sup> <i>ESI 17</i> 39	
5.2 Resultados obtidos na quantificação do DNA das amostras .....	40
5.3 Representação gráfica das frequências alélicas para 16 <i>loci</i> do <i>kit PowerPlex</i> <sup>®</sup> <i>ESI 17</i> .....	41
5.4 Resultados da análise de LD entre os 46 marcadores em estudo .....	44



## Índice de figuras

<b>Figura 1.1 Nacionalidades mais representativas no mapa dos estrangeiros residentes em Portugal.</b> O gráfico apresenta, em número e percentagem (parte superior), as principais nacionalidades estrangeiras em Portugal, descrevendo se houve um acréscimo ou decréscimo dessas populações (parte inferior), face ao ano de 2016. Fonte: (SEF, 2017). .....	3
<b>Figura 1.2 Mapa da América Latina. O Brasil, cuja capital federal é Brasília, divide a sua fronteira terrestre com todos os países da América do Sul, exceto Equador e Chile.</b> (Fonte: <a href="http://www.ensaiografico.com.br/wp-content/uploads/2017/02/mapa_america_sul.jpg">http://www.ensaiografico.com.br/wp-content/uploads/2017/02/mapa_america_sul.jpg</a> consultado a 18.01.2018). .....	4
<b>Figura 2.1 Fluxograma das várias etapas envolvidas no estudo de marcadores genéticos.</b> Representação esquemática das diversas etapas laboratoriais realizadas no estudo de marcadores do tipo InDel e marcadores do tipo STR, desde a colheita das amostras à análise estatística dos resultados obtidos. ....	11
<b>Figura 3.1 Exemplo de um eletroforegrama obtido aquando da deteção e análise dos marcadores do tipo InDel:</b> Genótipo de um indivíduo do sexo masculino oriundo do Brasil. Cada painel contém 7 a 8 marcadores, correspondendo a um máximo de 14 a 16 picos. ....	18
<b>Figura 3.2 Diferenças no tamanho (RFUs) dos picos do alelo correspondente à inserção do marcador HLD97 (D97+):</b> a) pode-se verificar a presença de um pico com 108 RFUs no local correspondente ao referido alelo e ausência de pico para o alelo de deleção D97-, indicando tratar-se de um indivíduo homozigótico para o locus em questão, apesar do tamanho reduzido do pico relativamente aos restantes; b) exemplo de um indivíduo homozigótico D97+, onde o alelo foi detetado de forma balanceada. ....	19
<b>Figura 3.3 Microvariantes detetadas aquando da análise dos marcadores InDel:</b> a) pico off ladder assinalado como D99+(-1), detetado no painel correspondente à cor verde, com um nucleótido a menos do que b) o alelo D99+ descrito; c) pico off ladder assinalado como D56-(+2), detetado no painel correspondente à cor azul, com dois nucleótidos a mais que d) o alelo D56- descritos. ....	20
<b>Figura 3.4 Microvariantes detetadas aquando da análise dos marcadores STR:</b> a) pico off ladder detetado no marcador SE33, utilizando o kit PowerPlex® ESI 17, confirmando-se, pela repetição da tipagem do indivíduo em questão com o kit GlobalFiler® b), que se trata do alelo microvariante 21.2; c) pico off ladder detetado no marcador D2S441 com o kit PowerPlex® ESI 17, comprovando-se a reprodutibilidade do alelo microvariante raro 14.1, pela repetição da tipagem do indivíduo, recorrendo aos kits c) GlobalFiler® e d) PowerPlex® Fusion 6C. ....	22
<b>Figura 3.5 Representação gráfica das frequências alélicas para os 30 marcadores InDel em estudo na população Brasileira residente em Lisboa.</b> ....	23
<b>Figura 3.6 Representação gráfica das heterozigotias observadas e esperadas para os 46 loci (30 InDel e 16 STR) estudados na população Brasileira residente em Lisboa.</b> ....	24

**Figura 3.7** **Árvore representativa das distâncias genéticas calculadas pelo índice de fixação. .... 27**

**Figura 5.1** **Representação gráfica das frequências alélicas para 16 loci do kit PowerPlex® ESI 17 (Promega) na população Brasileira residente em Lisboa. .... 41**

## Índice de tabelas

Tabela 2.1 Preparação das diluições em série a partir da solução Quantifiler THP DNA Standard de 100 ng/uL.....	12
Tabela 2.2 Componentes e respectivos volumes por reação de quantificação.....	13
Tabela 2.3 Componentes e respectivos volumes de preparação da master mix para amplificação com o Kit Investigator® DIPplex (QIAGEN).....	14
Tabela 2.4 Protocolo de amplificação do Kit Investigator® DIPplex (QIAGEN).....	14
Tabela 2.5 Componentes e respectivos volumes de preparação da master mix para amplificação com o kit PowerPlex® ESI 17 (Promega).....	15
Tabela 2.6 Componentes e respectivos volumes de preparação da master mix para amplificação com o kit PowerPlex® Fusion 6C (Promega).....	15
Tabela 2.7 Componentes e respectivos volumes de preparação da master mix para amplificação com o kit Globalfiler® (ThermoFisher) .....	15
Tabela 2.8 Reagentes necessários para a preparação do produto amplificado para eletroforese capilar e respectivos volumes, por cada reação. ....	16
Tabela 3.1 Dados relativos aos 46 loci estudados para a população migrante oriunda do Brasil residente na região de Lisboa: número de alelos estudados (2N), frequências alélicas, heterozigotia observada (Ho) e esperada (He), valor de p para o HWE. ....	24
Tabela 3.2 Matriz de distâncias genéticas entre a população em estudo - imigrantes Brasileiros residentes em Lisboa -, e entre seis populações anteriormente estudadas com os mesmos marcadores genéticos - Portugal e imigrantes de Angola, Moçambique, Guiné-Bissau, Cabo-verde e São Tomé e Príncipe. ....	26
Tabela 3.3 Parâmetros forenses calculados para os marcadores InDel (30 loci) e STR (16 loci) estudados na população migrante brasileira a residir na região de Lisboa.....	28
Tabela 5.1 Pannel de 30 marcadores, 29 loci InDel e a Amelogenina, incluídos no kit Investigator® DIPplex (QIAGEN) e pannel de 17 marcadores, 16 loci STR e Amelogenina, presentes no kit PowerPlex® ESI 17 (Promega). Para o primeiro kit, constam da tabela o nome dos marcadores, localização no cromossoma, sequência de inserção/deleção - HLD do inglês, Human Locus DIP – e respetiva referência na base de dados SNPforID. Adaptado de [64]. Para o segundo kit, constam o nome dos marcadores, localização no cromossoma, categoria e motivo de repetição e o N° - número - de alelos. Adaptado de [69]. ....	39

**Tabela 5.2 Resultados obtidos na quantificação do DNA das amostras estudadas.** Da tabela constam os códigos das amostras e respetivo resultado da quantificação do DNA em ng/μL. As amostras que, aquando da quantificação, não foi detetado DNA, apresentam-se com hífen. .... 40

**Tabela 5.3 Resultados obtidos aquando da análise de LD para os 46 loci em estudo:** pares de loci assinalados a cinzento, após correção Bonferroni, não se apresentam em LD ( $0,05 > p < 0,000048$ ), a cinzento escuro estão os pares de loci que, mesmo após a correção, mantêm-se em LD ( $p < 0,000048$ ). .... 44

## Lista de abreviaturas, siglas e unidades

A – Adenina

AIM – *Ancestral Informative Marker* – Marcador Informativo de Ancestralidade

CEE – Comunidade Económica Europeia

cM - Centimorgan

CODIS – *COmbined DNA Index System*

DIP – *Deletion/Insertion Polymorphism* – Polimorfismo de Deleção/Inserção

DNA – Ácido Desoxirribonucleico

dNTPs – Desoxinucleotídeos Trifosfatados

EDNAP – *European DNA Profiling*

ENFSI – *European Network of Forensic Sciences Institutes*

ESS – *European Standard Set*

Fst – Índice de Fixação

G – Guanina

He – Heterozigotia Esperada

HLD – *Human Locus DIP*

Ho – Heterozigotia Observada

HWE – Equilíbrio de Hardy-Weinberg

InDel – Inserção/Deleção

INMLCF – Instituto de Medicina Legal e Ciências Forenses

kb – quilo bases

LD – *Linkage Disequilibrium*

LE – *Linkage Equilibrium*

miRNA – MicroRNA

MP – *Matching Probability* – Probabilidade de Coincidência

Mpb – mega pares de bases

NJ – *Neighbor Joining*

*p* – Nível de Significância

PALOPs – Países Africanos de Língua Oficial Portuguesa

PCR – *Polymerase Chain Reaction* - Reação em Cadeia da Polimerase

PD – Poder de Discriminação

PE – Poder de Exclusão

PIC – *Polymorphism Information Content* – Conteúdo Informativo de Polimorfismo

RFUs – *Relative Fluorescence Units* – Unidades de Fluorescência Relativa

RNA – Ácido Ribonucleico

rpm – rotações por minuto

SGBF – Serviço de Genética e Biologia Forenses

SNPs – *Single Nucleotide Polymorphisms*

STRs – *Short Tandem Repeats*

TPI – *Typical Paternity Index* – Índice de Paternidade Típico

VNTRs – *Variable Number of Tandem Repeats*

# Capítulo 1 - Introdução

## 1.1 A geografia dos genes

Nos últimos 100.000 anos, o Homem anatomicamente moderno expandiu-se para ocupar toda a área habitável do globo [1]. O termo “anatomicamente moderno” é aplicado pela paleoantropologia, para distinguir a anatomia do *Homo sapiens*, consistente com a gama de fenótipos observada nos humanos contemporâneos, das variedades de humanos arcaicos extintos [2].

Ferramentas de disciplinas como arqueologia, fisiologia, antropologia e linguística, têm sido utilizadas para explorar a história desta expansão. Contudo, cada uma destas fontes tem as suas limitações, razão pela qual segundo Cavalli-Sforza, Menozzi e Piazza (1994, p. ix), para conclusões totalmente satisfatórias, é necessário comparar a informação extraída a partir destas áreas com evidências genéticas, visto os genes quase sempre apresentarem o nível de permanência necessário para discutir fissão, fusão e migração das populações, ao longo da história do Homem [1].

Antes da informação genética em larga escala estar disponível, foram avançadas duas hipóteses sobre o povoamento global, os modelos “*wave-of-advance*” e “*founder effect*”. O primeiro modelo - que defendia a substituição das populações -, foi inicialmente proposto para descrever a expansão da agricultura pela Europa. Nestes modelos, a transição neolítica foi acompanhada pela expansão de agricultores do Próximo-Oriente pela Europa, que total ou parcialmente substituíram caçadores-coletores residentes. Por outro lado, o modelo “*serial founder effect*”, propunha que as populações permaneceram nos locais que inicialmente haviam colonizado depois da expansão “*Out-Of-Africa*”, trocando migrantes lentamente com seus vizinhos mais próximos, até às migrações de longa distância há cerca de 500 anos atrás [1].

Segundo estudos mais recentes, após os avanços tecnológicos dos últimos anos, torna-se necessário reavaliar a história humana usando ferramentas de estudo do genoma modernas. De acordo com novos dados, grande parte da informação geográfica sobre as primeiras migrações humanas foi substituída por movimentos populacionais subsequentes, contrariando o segundo modelo, a favor da estabilização demográfica. A nova informação, indica ainda que, quando duas populações distintas se juntam durante a expansão demográfica, o resultado é *admixture* e não substituição completa da população, como defendia o primeiro modelo. Por esse motivo, K. Pickrell e Reich (2014) afirmam que, para uma melhor descrição dos aspetos importantes da história humana, é necessário elaborar novos tipos de modelos com *admixture* no seu centro [1].

## 1.2 Globalização e os fenómenos migratórios

A aplicação recorrente do termo globalização é relativamente recente - anos 90 -, erroneamente transmitindo a ideia de que se trata de um fenómeno totalmente novo. Na verdade, a sua origem remonta ao Século XIX, época em que muitos intelectuais recorreram à expressão, para salientar a importância da integração mundial aos vários níveis no processo de modernização. Segundo Alan J. Spector (2007), de uma forma geral, a globalização refere-se à integração mundial das pessoas, economia, cultura, tecnologia e, ainda, da comunicação e partilha de informação [3,4].

O desenvolvimento tecnológico representa um ponto catalisador para a globalização, facilitando não só a circulação de bens, serviços e pessoas, bem como, a comunicação entre as pessoas e instituições. Desta forma, não é de estranhar que se faça referência aos processos de globalização a propósito das migrações internacionais. O aumento na complexidade, volume e diversidade destes fenómenos, está fortemente associado aos avanços tecnológicos nas áreas dos transportes e das comunicações e, de um modo geral, aos processos de globalização [3,5].

A falta de oportunidades, o crescimento populacional, a opressão e os conflitos violentos nos países em desenvolvimento, bem como a crescente desigualdade doméstica e internacional, e a busca

persistente pela mão-de-obra qualificada e não qualificada, nos mercados de trabalho segmentado das sociedades ricas, são alguns dos fatores que contribuem para as migrações internacionais [5–7]. As tendências políticas com sentido à desregulamentação económica e liberalização dos mercados, que impulsionaram a globalização económica, são as mesmas que afetam fatores como a crescente segmentação do mercado de trabalho [5,8].

Neste sentido, se por um lado, a expansão da economia global providenciou ao Homem melhores oportunidades de vida, por outro, o impacto da globalização na desigualdade, verifica-se na crescente disparidade observada no padrão de vida e no nível de segurança humano disponível para pessoas em diferentes partes do mundo [9].

Uma marcante consequência da intensificação nas discrepâncias, fruto da crescente globalização, é o aumento em escala e extensão das migrações internacionais. Globalmente, o número de migrantes internacionais em 2015 excedia os 240 milhões, cerca de 60% mais do que em 1990 [5,9,10].

### **1.3 Imigração Brasileira em Portugal**

Portugal, considerado país de emigração até décadas mais recentes – anos 70 e 80 –, abraçou uma nova realidade, recebendo imigrantes do Brasil, África, Europa Central e de Leste [11,12]. Por outro lado, durante esse mesmo período, o Brasil que desde a sua descoberta até ao final da Segunda Guerra Mundial acolhera imigrantes, devido á crise social e económica que se havia instalado, viu-se na necessidade de recorrer à emigração [11].

Nas últimas quatro décadas, Portugal foi palco de dois principais fluxos migratórios provenientes do Brasil. O primeiro em 1980, época em que após a sua democratização, Portugal procurava por mão-de-obra qualificada, recebendo migrantes Brasileiros em busca de novas oportunidades, num país onde as barreiras linguísticas e culturais fossem menores. Marcado pelo aumento na migração ilegal, o segundo fluxo surge em 1986, após a entrada de Portugal na Comunidade Económica Europeia (CEE).

Em 2000, por ocasião das comemorações dos quinhentos anos do descobrimento do Brasil, foi assinado o celebre “Tratado de Amizade, Cooperação e Consulta”, onde Portugueses e Brasileiros residentes, beneficiários do Estatuto de Igualdade, disfrutaram de um conjunto mais alargado de direitos civis e políticos, estando sujeitos aos mesmo deveres dos nacionais. Todavia, o tratado nada veio a resolver quanto ao aumento na imigração ilegal entre os países, questão abordada pouco tempo depois – entre 2003 e 2004 -, com celebração de novo acordo bilateral – conhecido como “Acordo Brasil-Portugal Sobre a Contratação Recíproca de Nacionais” [11,13,14].

Com efeito, a escalada no número de imigrantes Brasileiros residentes em Portugal, fez-se sentir em 2007, quando o número de migrantes com esta nacionalidade superou os Cabo-verdianos, mantendo-se, desde então, como a principal comunidade estrangeiro residente em Portugal [11].

De acordo com o Instituto Nacional de Estatística, entre 1990 e 2010, verificou-se um aumento no número de imigrantes Brasileiros em Portugal, chegando a totalizar cerca de 119 000 indivíduos, no fim desse período. Por outro lado, entre os anos 2011 e 2016, o número de residentes com esta nacionalidade veio tendencialmente a cair, apresentando-se em território nacional, cerca 80 000 indivíduos em 2016, uma diminuição de 31.9% face ao ano de 2010 [15].

O agravamento da austeridade em Portugal e o aumento das aquisições de nacionalidade Portuguesa – sensivelmente 10 000 pareceres positivos em 2016 –, são dois fatores que terão contribuído para o decréscimo na população imigrante Brasileira [16].

O relatório de imigração, fronteiras e asilo, referente ao ano de 2017, revelou uma reversão na diminuição do número de imigrantes com esta nacionalidade, verificando-se um aumento de 5,6 % face a 2016 (figura 1.1) [17]. Segundo Pereira e Esteves (2017), a ténue melhoria dos indicadores económicos em Portugal e a mais recente crise política e económica instalada no Brasil, parecem estar a fomentar



novas ondas migratórias. Adicionalmente, apesar da conjuntura económica, a facilidade linguística, a conservação de certos nichos de empregabilidade – nomeadamente na estética, comida pronta e doçaria – e a qualidade de vida experimentada em Portugal, são favoráveis à manutenção da imigração Brasileira em Portugal [18].

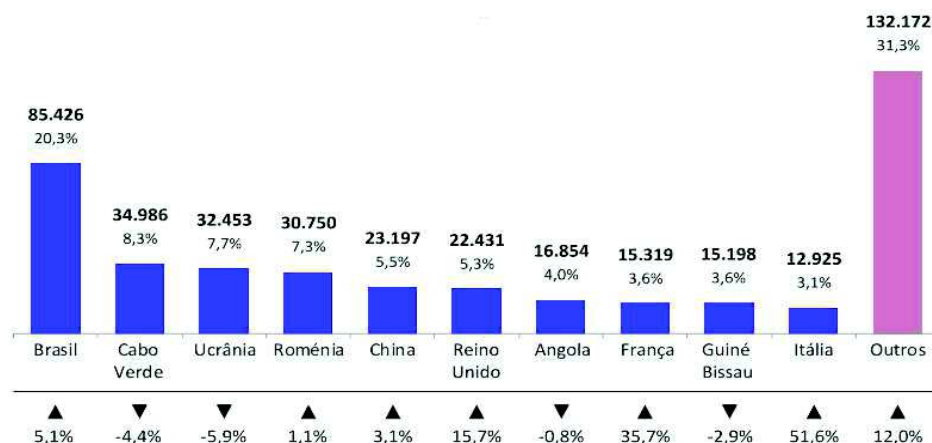


Figura 1.1 **Nacionalidades mais representativas no mapa dos estrangeiros residentes em Portugal.** O gráfico apresenta, em número e percentagem (parte superior), as principais nacionalidades estrangeiras em Portugal, descrevendo se houve um acréscimo ou decréscimo dessas populações (parte inferior), face ao ano de 2016. Fonte: (SEF, 2017).

### 1.3.1 República Federativa do Brasil

Um dos episódios que marcou a expansão marítima portuguesa, iniciada no princípio do século XV, foi a descoberta do Brasil por Pedro Álvares Cabral em abril de 1500 [19]. O Brasil, oficialmente conhecido como República Federativa do Brasil, a partir da constituição de 1967 durante o período de Ditadura Militar (1964-1985), é o quinto maior país do mundo em termos populacionais - aproximadamente 207 000 000 habitantes em 2017 - e de extensão territorial, com cerca de 8 515 760 km<sup>2</sup>. Ocupando praticamente metade do continente da América do Sul, o Brasil encontra-se delimitado a leste pelo Oceano Atlântico, ao longo dos sensivelmente 7400 km<sup>2</sup> que compõem a sua costa litoral [20]. O Brasil divide limites terrestres com dez países da América do Sul - Uruguai, Argentina e Paraguai, Bolívia, Peru e Colômbia, Venezuela, Suriname, Guiana e ainda, com o Departamento Ultramarino Francês da Guiana -, numa extensão que chega aos 15 700 km<sup>2</sup> (Figura 1.2) [21].

A população Brasileira, altamente miscigenada – principalmente por Europeus, Africanos e Nativo-Americanos -, é um reflexo das sucessivas ondas migratórias que deram lugar à sua heterogeneidade [22]. A população Portuguesa foi a primeira a desembarcar no Brasil, durante o período de colonização Europeia – entre os anos 1500 e 1800 -, trazendo consigo cerca de 500 000 colonizadores, na sua maioria homens, que contactaram com ameríndios e sensivelmente 4 000 000 escravos da África Subsaariana. A alteração demográfica das Américas, após a chegada dos Europeus e Africanos, foi de tal forma significativa, que grande parte da ascendência dos seus atuais habitantes não deriva dos Nativo-Americanos – analogamente ao que ocorre noutras partes do globo [1,23].

Aquando da invasão do Brasil pelos Portugueses, estima-se que existissem cerca de 1000 povos indígenas, o que corresponderia a uma população em torno dos 2 a 4 milhões de habitantes. Desde então e até à contemporaneidade, o genocídio, em comunhão com a escravidão e as epidemias, são uma realidade que continuamente tem assolado e dizimado estes povos. Segundo o último censo demográfico do Brasil, conduzido em 2010, as raças, branca, parda e preta, eram as mais frequentes. Nesse mesmo censo, estimou-se que a população Brasileira chegasse aos 190 000 000 habitantes [20,24–29].



Figura 1.2 Mapa da América Latina. O Brasil, cuja capital federal é Brasília, divide a sua fronteira terrestre com todos os países da América do Sul, exceto Equador e Chile. (Fonte: [http://www.ensaiografico.com.br/wp-content/uploads/2017/02/mapa\\_america\\_sul.jpg](http://www.ensaiografico.com.br/wp-content/uploads/2017/02/mapa_america_sul.jpg) consultado a 18.01.2018).

A partir do início do século XVIII e até 1975, com a emergência de uma nova onda migratória, chegam ao Brasil aproximadamente 6 000 000 migrantes oficiais provindos de diferentes países – sobretudo Itália, Portugal, Espanha e Alemanha, Síria, Líbano e Japão -, acrescentando complexidade à já multiétnica população Brasileira [30].

Um aspeto bem assente sobre estas migrações, é precisamente o seu efeito na estrutura das populações, representando um dos principais fatores de variabilidade genética a nível intrapopulacional e interpopulacional.

#### 1.4 Genética Populacional

A genética populacional, é a ciência vocacionada para o estudo da distribuição e alteração das frequências alélicas e genotípicas, ao nível intrapopulacional e interpopulacional ao longo das gerações, bem como dos fatores evolutivos – genéticos, sociais e ambientais – promotores de variação genética das populações. Importará ainda referir, que em genética populacional, uma população define-se como um grupo de indivíduos que ocupam igual tempo e espaço, sendo capazes de se cruzarem entre si [31,32]. É na definição do Princípio de Hardy-Weinberg, demonstrado independentemente por G.H. Hardy e Wilhelm Weinberg, que se encontram os pilares fundadores da genética de populações, cujo seu atual reconhecimento deve-se em grande parte ao trabalho de autores como J. B. S. Haldane, R. A. Fisher e Sewall Wright [33].

O estudo genético das populações desenvolvido pela genética populacional, constitui uma ferramenta especialmente útil no contexto da genética forense, permitindo, pela estimativa das frequências e combinações alélicas de grupos específicos de indivíduos, construir bases de dados populacionais, aplicando-as para inferir acerca da raridade de certo perfil genético. É a partir da contagem do número de vezes que cada alelo é observado, que as frequências alélicas são aferidas para a população de interesse, assim, quanto maior o número de indivíduos que a compõe, tanto mais representativa e precisa será a base de dados [34].

Para garantir que as frequências alélicas obtidas, respeitam as transposições da genética mendeliana para a população, testes estatísticos são tipicamente aplicados para avaliar a independência dos alelos

num *locus* – teste do Equilíbrio de Hardy-Weinberg – e entre *loci* – *Linkage Equilibrium*. Só assim, podem os dados ser utilizados para estimar a raridade do genótipo, a partir do cálculo da probabilidade de coincidência [34,35].

#### 1.4.1 Equilíbrio de Hardy-Weinberg

Para o estudo dos fenómenos populacionais, torna-se mais útil recorrer a frequências alélicas, uma vez que ao contrário dos genótipos, cujas frequências se alteram em resposta a processos de segregação e recombinação, alelos raramente sofrem mutação numa determinada geração, sendo mais estável a sua transmissão ao longo das gerações. Invariavelmente, frequências alélicas podem ser calculadas a partir das frequências genotípicas, contudo, para que o inverso se verifique, é imperativo que a variação genética numa população permaneça constante entre gerações.

A lei do equilíbrio de Hardy-Weinberg – HWE, do inglês *Hardy-Weinberg equilibrium* – afirma que, uma população panmítica, indefinidamente grande, em que a migração, mutação e seleção natural são insignificantes e as frequências alélicas iguais entre machos e fêmeas, atinge o equilíbrio genético ao fim de uma geração [36,37].

Se dois alelos *A* e *a* ocorrerem na população com frequências *p* e *q*, respetivamente, o genótipo AA (homozigoto) deverá ocorrer com frequência  $p^2$  e o genótipo Aa (heterozigoto) a uma frequência  $2pq$ . As frequências genotípicas esperadas – calculadas a partir das alélicas –, são então comparadas às frequências genotípicas observadas e, caso os valores sejam semelhantes, assume-se que os alelos dentro do *locus* genético estão em equilíbrio [34,38].

Teoricamente, sobre estas condições, as frequências dever-se-iam manter constantes durante muitas gerações. Porém, fatores como seleção natural, mutação, endogamia e subestrutura populacional, presentes nas populações naturais, contribuem para desvios do HWE, pois violam os pressupostos necessários à manutenção da estabilidade genética [34,39]. Uma vez que, pequenos desvios não condicionam a utilização de uma base de dados, nesses casos, alguns autores farão pouco mais do que observar que desvios estatisticamente significativos do HWE foram registados, para certo *locus* do conjunto de dados da população estudada [34].

Outro motivo para a aplicação do teste do HWE, consiste em procurar qualquer indício de excesso de homozigoticidade, devido, essencialmente, à presença de alelos nulos resultantes de *drop-out* alélico, através da comparação do número observado e esperado de homozigotos [40].

#### 1.4.2 *Linkage Dis(equilibrium)*

O conceito de *linkage*, surge em 1905, pelo estudo da cor da flor e forma do pólen em *Lathyrus odoratus*, onde os geneticistas Reginald Punnett, William Bateson e Edith Rebecca Saunders, observaram uma tendência para a hereditariedade conjunta destas características, constituindo uma das primeiras exceções às leis de Mendel. Subsequentemente, em 1910/11, Thomas H. Morgan, demonstrou que o nível de associação dos genes no cromossoma X de *Drosophila melanogaster*, variava entre os diferentes pares [41–43].

O *linkage* genético refere-se à tendência de dois ou mais genes, dispostos linearmente no mesmo cromossoma, durante o processo biológico de hereditariedade serem transmitidos em bloco, consoante o seu nível de recombinação meiótica. Quanto maior a proximidade entre dois ou mais *loci*, menor a taxa de recombinação genética entre eles, aumentando a probabilidade de serem transmitidos em conjunto. Por outro lado, mesmo que em associação, *loci* que se encontram a 50 cM de distância, recombinaem com a mesma frequência com que permanecem juntos, comportando-se como genes não associados que segregam independentemente [34,43–45].

Os termos *linkage equilibrium* (LE) e *linkage disequilibrium* (LD), em semelhança ao conceito de ligação genética, surgem para exprimir a probabilidade de alelos em diferentes *loci* serem co-herdados.

Assim, diz-se que estão em LE, alelos em associação aleatória, onde o genótipo encontrado em certo *locus* é independente do genótipo de um *locus* distinto – por exemplo, genes em cromossomos diferentes. Em contrapartida, a associação não aleatória de alelos em dois ou mais *loci* – presentes ou não no mesmo cromossoma – poderá corresponder a LD, isto é, ocorre quando dois genótipos em determinados *loci* não são independentes um do outro, medindo o grau em que os alelos nesses *loci* se encontram associados [34,46].

O LD é importante para a biologia evolutiva e genética humana, porque afeta e é afetado por diversos fatores, tais como mutações recentes, deriva genética, fluxo de genes, consanguinidade e especialmente pela seleção natural. Em genética forense, a independência dos *loci* é essencial para o cálculo da probabilidade de coincidência, ou frequência de determinado perfil, obtida a partir da multiplicação das frequências genótípicas em *loci* individuais [47,48].

## 1.5 Genética Forense

À genética forense compete, a análise comparativa dos perfis genéticos obtidos a partir do estudo de amostras biológicas ou outras, sejam amostras de referência ou problema, prestando deste modo auxílio aos processos judiciais e investigação criminal [49].

O perfil genético compreende o conjunto de informação relativa à constituição genética de um indivíduo num número mínimo de *loci*, com vista à sua identificação, sendo que o tipo e número de marcadores genéticos a utilizar, está em conformidade com certos requisitos legais e científicos. A informação genética, pode apresentar-se sobre forma de uma sequência específica de nucleótidos – haplótipo de DNA mitocondrial – ou numericamente, correspondendo ao número de repetições da unidade do bloco de repetições – como num perfil genético de STRs – do inglês, *Short Tandem Repeats* – autossômicos [49–52].

Em 1901, quando Karl Landsteiner descobriu o sistema sanguíneo ABO, apercebeu-se do potencial forense desta variação, que se transformou numa das primeiras evidências genéticas usadas em tribunal. Apesar deste e outros sistemas sanguíneos, terem a capacidade de excluir certos indivíduos, a partir do resultado de testes serológicos baseados em anticorpos, provar a origem da amostra biológica é mais difícil e depende da variabilidade do sistema de tipagem. Até à década de 1980, métodos serológicos e eletroforéticos de proteínas eram utilizados para o estudo da variabilidade dos grupos sanguíneos e proteínas polimórficas. Contudo, para além de apresentarem um nível de variabilidade relativamente baixo, estes marcadores tinham a tendência para rapidamente se degradarem ou eram comprometidos por enzimas bacterianas [50].

A técnica de “*DNA Fingerprinting*”, atualmente conhecida como tipagem de DNA, foi inicialmente descrita pelo geneticista Alec Jeffreys em 1985, após identificar regiões polimórficas do DNA que denominou VNTRs – do inglês, *Variable Number of Tandem Repeats*. A aplicabilidade forense desta técnica, tornou-se evidente em 1986, quando Alec Jeffreys, a utilizou, a pedido da polícia inglesa, para resolver um caso de dupla violação e homicídio. Surpreendentemente, o perfil genético obtido a partir dos vestígios de sêmen encontrados em cada local de crime, não coincidia com o perfil do suspeito que havia confessado tê-los cometido, mas sim com Colin Pitchfork, que persuadira o amigo, Richard Buckland, a entregar uma amostra em seu lugar. Confrontado com as evidências, o verdadeiro culpado confessou a autoria dos crimes, evidenciando a dimensão política e social da técnica que veio a revolucionar a investigação criminal [34,53].

Os avanços na tecnologia do DNA, ao longo das últimas três décadas, têm contribuído para o desenvolvimento de técnicas de identificação forense extraordinariamente precisas. Porém, certas aplicações têm desencadeado controvérsias acerca da privacidade genética, principalmente após a construção de bases de dados de DNA com fins forenses [54].

Atualmente, cabe aos Serviços de Biologia e Genética Forenses – SGBF –, do Instituto Nacional de Medicina Legal e Ciências Forenses – INMLCF, I.P. –, a realização de perícias e exames de identificação genética, nomeadamente os de investigação de parentesco biológico, identificação individual e criminalística biológica, tais como crimes sexuais ou homicídios.

## **1.6 Marcadores moleculares em genética forense**

A obtenção de perfis genéticos destinados à identificação humana, requer o estudo de marcadores moleculares, genes ou sequências de DNA com localização conhecida nos cromossomas. Uma vez que, 99,7 % do genoma humana não varia na população, os marcadores que geneticamente permitem diferenciar indivíduos estão localizados nos restantes 0,3 % [34,49].

Consoante o tipo de marcador molecular usado, o perfil genético obtido, pode ou não ser individualizante, isto é, poderá identificar apenas uma pessoa, sendo-lhe exclusivo ou, por outro lado, identificar um conjunto de pessoas [49].

Os principais tipos de marcadores autossómicos, que permitem a definição de perfis genéticos individualizantes, são os STRs e InDels – inserções/deleções –, abordados separadamente nos subcapítulos que se seguem e, ainda, os SNPs – do inglês, *single nucleotide polymorphisms* [50]. Os perfis genéticos específicos de uma linha paterna ou materna e, portanto, não individualizantes, são obtidos a partir da análise de marcadores do cromossoma Y ou cromossoma X, respetivamente. Adicionalmente, e porque apresenta hereditariedade exclusivamente materna, o DNA mitocondrial também se apresenta como marcador de linhagem, refletindo a ascendência matrilinea de um indivíduo [50,54].

Estudos mais recentes, têm revelado as vantagens e limitações da aplicação forense de miRNAs tecido-específicos, como marcadores para determinação do tipo de fluidos corporais presentes em manchas colhidas nos locais de crime [55,56].

### **1.6.1 Marcadores do tipo InDel**

Os marcadores moleculares baseados em polimorfismos de inserção ou deleção, são marcadores bialélicos que se caracterizam pela presença ou ausência de sequências específicas de nucleótidos, cujo tamanho varia, maioritariamente, entre 2 a 16 pb, com comprimento máximo de 55 pb. Os InDels, são a classe mais abundante de polimorfismos a seguir aos SNPs, representando cerca de 30% de toda a variação presente no genoma humano, apresentando-se com uma densidade média de 1 por cada 7,2 kb de DNA [57,58].

Os marcadores do tipo InDel constituem uma mais-valia no âmbito da identificação genética para fins forenses, como método complementar aos STRs, especialmente em amostras biológicas muito degradadas, devido ao pequeno tamanho dos fragmentos de amplificação (até 160 pb), para além de apresentarem taxas de mutação baixas –  $2,0 \times 10^{-9}$  –, comparativamente a outros marcadores moleculares, sendo, por isso, também vantajosos em casos complexos de filiação biológica [59,60].

Analogamente aos InDel, os marcadores do tipo SNP, para além de não produzirem picos *stutter*, apresentam produtos de amplificação de tamanho reduzido e taxa de mutação inferior à dos STRs. Contudo, a obtenção de um poder de discriminação semelhante a um painel com 10 a 16 *loci* STR, requer a análise de 50 a 100 SNPs, que exigem técnicas de deteção mais complexas. Uma das desvantagens dos marcadores InDel, refere-se também ao seu poder de discriminação, que uma vez inferior ao dos microssatélite, requer o estudo de 30 a 40 *loci*, para alcançar uma capacidade de diferenciação idêntica ao painel de marcadores do CODIS – COmbined DNA Index System. Por outro lado, a análise de marcadores do tipo InDel é em muito semelhante à dos STRs e, portanto, relativamente simples e automatizada, bem como menos dispendiosa que a dos SNPs [59].



Porque combinam características desejáveis de STRs e SNPs, os polimorfismos InDel apresentam particularidades interessantes como marcadores genéticos, com considerável potencial em contexto forense, inclusive em estudos de ancestralidade, onde são aplicados como marcadores informativos de ancestralidade – AIMS –, visto divergirem significativamente em termos de frequências alélicas, entre grupos populacionais geograficamente afastados [61,62].

Porque se apresentam distribuídos por todo o genoma humano, os InDels são facilmente genotipados, permitindo, na construção de painéis de estudo, uma variedade de combinações de marcadores [63].

Presentemente, são conhecidas cinco classes de InDel: inserções e deleções de um par de bases; expansões monoméricas ou expansões de unidades de repetição com 2 a 15 pb; inserções de transposições ou sequências aleatórias de DNA, na sua maioria de tamanho inferior a 100 pb, que representam cerca de 41% dos InDel identificados [58].

#### **1.6.1.1 – Sistema *Investigator*<sup>®</sup> *DIPplex***

Presentemente, existem pelo menos dois *kits* comercialmente disponíveis para amplificação simultânea de diversos marcadores InDel, desenvolvidos para identificação humana. Os marcadores incluídos nesse tipo de sistemas *multiplex*, devem obedecer a critérios específicos, tais como, serem bialélicos e não codificantes, apresentarem frequências alélicas mínimas maiores que 0.25 nas populações Europeia, Africana e Asiática, apresentarem uma heterozigocidade média maior que 0.40 e uma variação dos alelos entre 2 pb e 5 pb [61].

Para o presente trabalho, utilizou-se o *kit Investigator*<sup>®</sup> *DIPplex* (*QIAGEN*), que permite a tipagem 30 *loci*, incluindo a amelogenina, numa única reação *multiplex*. Estes marcadores encontram-se distribuídos por 19 autossomas, a uma distância mínima de 10 Mpb – mega pares de base –, e o gene da amelogenina nos cromossomas X e Y, consoante se trate, respetivamente, da AMELX ou AMELY. Os marcadores encontram-se sinalizados pela marcação fluorescente dos *primers* utilizados, pelos fluorocromos 6-FAM – azul –, BTG – verde –, BTY – amarelo – e BTR – vermelho, facilitando a análise. As características destes marcadores encontram-se indicadas na tabela 5.1 dos anexos [64].

Apesar do *kit Investigator*<sup>®</sup> *DIPplex* já ter sido aplicado em vários estudos das populações humanas da Europa, Ásia, África e América, poucos foram os realizados nas populações da América do Sul [65].

#### **1.6.2 Marcadores do tipo STR**

Os STRs fluorescentes, descritos pela primeira vez em 1991, substituíram os VNTRs, que desde 1989, devido a questões relativas ao uso de estatísticas para a genética populacional e à qualidade dos resultados obtidos em laboratórios forenses, impuseram-se desafios à sua admissibilidade [34].

Também conhecidos como microssatélites, os STRs representam cerca de 3% do genoma humana, estimando-se que, em média, ocorram a cada 100 000 nucleótidos. Estas regiões do DNA, são compostas por repetições perfeitas ou quase perfeitas – nesse caso, designadas microvariantes –, em que as unidades de repetição com 2 pb a 7 pb de comprimento, repetem-se justapostamente de 2 a 40 vezes. Consoante o tamanho, as unidades de repetição podem classificar-se como mono-, di-, tri-, tetra-, penta- ou hexanucleotídicas, sendo que as repetições tetranucleotídicas são as mais utilizadas em identificação humana, visto serem mais estáveis durante a amplificação. Ao longo da reação de PCR – do inglês, *polymerase Chain Reaction* –, podem formar-se artefactos resultantes da amplificação dos microssatélites, que surgem nos eletroforegramas como picos *stutter* com altura inferior a cerca de 15% do alelo original, no caso dos STRs formados por sequências repetitivas de quatro nucleótidos. Tratando-se de STRs di- ou trinucleotídicos, os mesmos podem apresentar picos de *stutter* com cerca de 30% ou mais da altura do alelo verdadeiro, dificultando a leitura de misturas. Dada a raridade dos STRs penta- e hexanucleotídicos no genoma humano, é pouco frequente a sua utilização em genética. No que

concerne ao padrão de repetição dos microssatélites, estes podem ainda dividir-se em simples, compostos e complexos, consoante o número de unidades de repetição, bem como do seu comprimento e sequência [34,66].

Apesar do DNA em algumas amostras forenses, apresentar-se em quantidades vestigiais ou com certo nível de degradação, o pequeno tamanho dos alelos STR –  $\approx 100$  pb a 400 pb –, permite-lhes serem analisados após amplificação pela reação de polimerização em cadeia. Esta metodologia de tipagem do DNA, que recorre a técnicas mais sensíveis, automatizadas e que requerem menor dispêndio de tempo, superou as limitações associadas à manipulação de sequências VNTR, o que contribuiu para a utilização dos microssatélites como prática generalizada nos serviços e laboratórios de genética forense. Mais ainda, estes marcadores apresentam um elevado poder de discriminação, sensibilidade e habilidade para resolver misturas, como as obtidas em casos de agressão sexual, contendo material biológico da vítima e agressor [34,50,67].

#### 1.6.2.1 – Sistema PowerPlex® ESI 17

Dos inúmeros kits comercialmente disponíveis para amplificação simultânea de vários *loci* STR, utilizou-se, para o presente estudo, o Sistema *PowerPlex® ESI 17*, desenvolvido para identificação humana, em casos de investigação de filiação ou criminalística biológica, após a última expansão do painel de marcadores forenses do ESS, que ocorreu em 2009. Tucker e colaboradores, em 2011, destacaram os bons resultados obtidos com o kit, não só em amostras com presença de inibidores, mas também em amostras de DNA fragmentado, por incluir mini-STRs [68].

O referido kit permite a co-amplificação e deteção baseada em fluorescência de 17 *loci*, pela marcação dos *primers* com diferentes fluorocromos. Os *primers* dos *loci* D22S1045, D2S1338, D19S433, D3S1358 e o gene homólogo da amelogenina – cromossoma X e Y –, encontram-se marcados com fluoresceína, que os sinaliza com a cor azul. Os *loci* D2S441, D10S1248, D1S1656, D18S51 e D16S539, têm um dos seus *primers* marcados com o fluoróforo JOE, que emite fluorescência na região do verde. O marcador TMR-ET, sinaliza com a cor amarela, que no eletroforegrama corresponde ao painel preto, os *loci* D12S391, D21S11, vWa e TH01, ao passo que o CXR-ET, sinaliza com a cor vermelha, os *loci* SE33, FGA e D8S1179 [69].

Os *loci* amplificados pelo kit, apresentam fragmentos de amplificação de tamanho reduzido, variando entre 72pb – TH01 – e 345pb – D22S1045, como recomendado pela comunidade forense Europeia. Adicionalmente, todos os *loci* obedecem à necessidade dos marcadores se encontrarem, preferencialmente, em cromossomas distintos, à parte dos *loci* D2S1338 e D2S441, ambos no cromossoma 2 e dos *loci* D12S391 e vWa, no cromossoma 12. Ainda assim, cada um desses marcadores encontra-se em braços diferentes do cromossoma em que se situam [69,70]. As características dos 16 *loci* STR e Amelogenina, incluídos no Sistema *PowerPlex® ESI 17*, encontram-se na tabela 5.1 na seção dos anexos.

#### 1.6.3 European Standard Set (ESS)

Para facilitar o intercâmbio de perfis das bases de dados de DNA entre laboratórios, a comunidade Europeia, em 1999, através do trabalho colaborativo entre o *European Network of Forensic Science Institutes* – ENFSI – e o *European DNA Profiling* – EDNAP –, decidiu definir um conjunto mínimo de *loci*. Naquela época, o painel de marcadores genéticos incluía sete *loci* – TH01, vWA e FGA, D8S1179, D18S51, D21S11 e D3S1358 –, comuns a todos os laboratórios Europeus. Em 2009, cinco novos mini-STR – D12S391 e D1S1656, D2S441, D10S1248 e D22S1045 – foram adicionados ao painel, com o objetivo de aumentar o poder de discriminação e melhorar a robustez e qualidade dos resultados, bem como a sensibilidade do teste, para que quantidades menores de DNA sejam detetadas [59,70,71].

A constituição de um painel de marcadores para identificação humana, requer a seleção de *loci* que cumpram com determinados requisitos, nomeadamente, elevado poder de discriminação, com uma heterozigotia observada de 70%, reduzido tamanho dos fragmentos de amplificação, localização em cromossomas distintos ou com localização afastada em relação a outros *loci*, robustez e reprodutibilidade dos resultados e, ainda, baixa taxa de mutação [71].

## 1.7 Objetivos

Os imigrantes Brasileiros, desde 2007, representam a principal comunidade estrangeira residente em Portugal, totalizando, em 2017, cerca de 85 000 indivíduos, na sua maioria registados no distrito de Lisboa. Presentemente, não existem dados disponíveis para os marcadores InDel na população Brasileira residente em Lisboa, nem para os últimos marcadores STR adicionados ao sistema ESS.

Assim, pretende-se com o presente estudo, caracterizar geneticamente a população de imigrantes do Brasil em Lisboa, com marcadores do tipo InDel e STRs do último painel Europeu de marcadores genéticos, para que possam ser utilizados nos casos periciais de genética forense, que envolvam indivíduos da referida população.

Especificamente, pretende-se com o estudo:

- Caracterizar indivíduos imigrantes do Brasil com PCR *multiplex* de marcadores do tipo InDel, capazes de definir perfis individuais, e STRs do último painel Europeu de marcadores moleculares;
- Calcular as frequências alélicas dos referidos marcadores na população em estudo;
- Calcular outros parâmetros populacionais e de interesse forense dos *loci* estudados na população imigrante Brasileira residente em Lisboa, nomeadamente, HWE, heterozigotia observada ( $H_o$ ) e heterozigotia esperada ( $H_e$ ), probabilidade de coincidência ou probabilidade de *matching* (MP), poder de discriminação (PD) e poder de exclusão (PE), conteúdo informativo de polimorfismo (PIC) e índice de paternidade associado (TPI);
- Comparar as frequências alélicas obtidas para a população imigrante Brasileira com a população portuguesa, e com as populações imigrantes PALOPs – Angola, Guiné-Bissau, Moçambique, São Tomé e Príncipe e Cabo Verde (PALOPs) –, anteriormente estudadas pelo



nosso grupo, utilizando os mesmos marcadores, inferindo acerca da diferenciação genética ( $F_{st}$ ) entre as mesmas.

## Capítulo 2 – Materiais e Métodos

### 2.1 Procedimento experimental

A elaboração do presente estudo, requereu a execução de diversas técnicas laboratoriais para cada um dos tipos de marcadores utilizados, estando todas essas etapas identificadas no fluxograma que se segue (Figura 2.1).

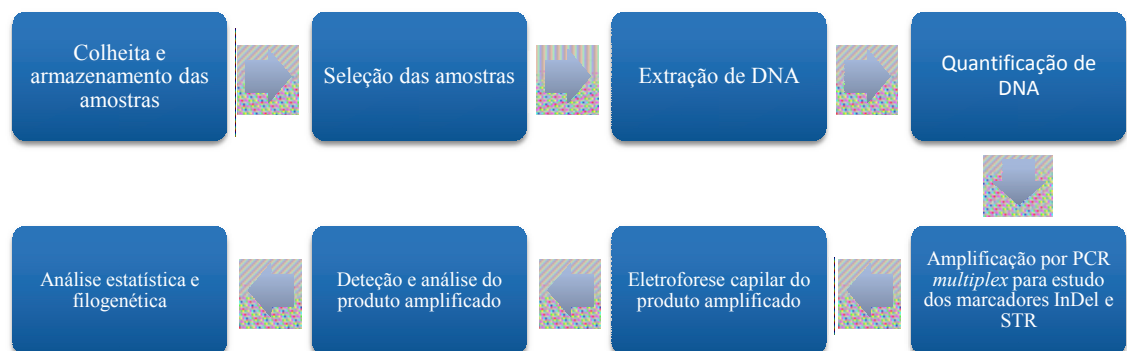


Figura 2.1 **Fluxograma das várias etapas envolvidas no estudo de marcadores genéticos.** Representação esquemática das diversas etapas laboratoriais realizadas no estudo de marcadores do tipo InDel e marcadores do tipo STR, desde a colheita das amostras à análise estatística dos resultados obtidos.

### 2.2 Colheita e armazenamento das amostras

As manchas de sangue utilizadas no presente trabalho, foram colhidas em cartões Whatman™ (GE Healthcare) no SGBF do INMLCF, I.P. As amostras, uma vez colhidas, foram colocadas na estufa para secagem e armazenadas em envelopes de papel à temperatura ambiente.

### 2.3 Amostragem

Para o presente trabalho, foram estudados 244 indivíduos imigrantes residentes em Lisboa, com naturalidade e ancestralidade Brasileira, não relacionados entre si, os quais compareceram ao SGBF da

Delegação do Sul do INMLCF, I.P., no âmbito de ações de investigação de parentesco biológico entre os anos 2002 e 2016.

## 2.4 Extração de DNA

A extração de DNA das amostras foi concretizada pelo método de Chelex 100® (BioRad), descrito em 1991, por Walsh e seus colaboradores [72]. O Chelex é uma resina quelante com elevada afinidade para iões metálicos polivalentes, tais como o magnésio, cofator de nucleases responsáveis pela degradação do DNA. Este método de extração apresenta diversas vantagens, tais como simplicidade e custo reduzido, para além dos reagentes utilizados serem inócuos à saúde. O Chelex 100® (BioRad) a 5% foi preparado previamente à extração, dissolvendo 2,5 g deste pó em 50 mL de água desionizada.

O procedimento de extração de DNA por este método, iniciou-se com a preparação de um fragmento circular do papel de suporte da mancha de sangue para um tubo de 1,5 mL, identificado com o código atribuído à amostra, com o auxílio de um *paper-punch single hole* de 3mm de diâmetro. De modo a evitar contaminação entre amostras, sempre que é efetuado um corte com o alicate *paper-punch single hole*, este deverá ser limpo com álcool a 70%. Seguidamente, adicionou-se ao tubo 1 mL de água desionizada, deixando incubar à temperatura ambiente durante cerca de 20 minutos, para que todo o material passe para a solução. Decorrido esse tempo, agitaram-se os tubos no vórtex durante 10 segundos, a baixa rotação, com posterior centrifugação a 14 000 rpm (*Centrifuge 5804 R, Eppendorf*), durante 3 minutos, a uma temperatura de 20°C. Com auxílio de uma micropipeta, retirou-se cerca de 950 µL de sobrenadante, sem perturbar o *pellet*, e adicionou-se 180 µL de Chelex 100® (BioRad) a 5%, tendo em atenção que as partículas se encontrem em suspensão após agitação (manual ou automática), de modo a serem aspiradas de forma homogênea. Ulteriormente, colocaram-se os tubos num banho a 56°C durante cerca de 20 minutos. Findo esse tempo, as amostras foram novamente agitadas no vórtex durante 15 segundos, a alta velocidade, colocando-as em suportes flutuantes para o banho de ebulição a 100°C, durante 8 minutos. Esta temperatura promove a rutura das membranas e destruição das proteínas celulares, libertando-se assim as moléculas de DNA para a solução. Por fim, os tubos foram agitados no vórtex, a alta velocidade, durante alguns segundos, e centrifugados a 14 000 rpm, durante 5 minutos, a 20°C. O DNA extraído foi armazenado a -20°C até ser necessário para a PCR a Tempo Real e *multiplex*.

## 2.5 Quantificação de DNA

A quantificação de DNA foi realizada recorrendo ao *Kit Quantifiler™ Trio*, de acordo com as condições do fabricante, descritas no manual de utilização. Este *Kit* é composto por todos os reagentes necessários à amplificação, deteção e quantificação do DNA das amostras em estudo. Nele está incluído o *Quantifiler™ Trio Primer Mix*, com os pares de *primers* para as três regiões amplificadas e respetivas sondas, o *Quantifiler™ THP PCR Reaction Mix*, com DNA polimerase, dNTPs e tampão e, ainda, o *Quantifiler™ THP DNA Dilution Buffer* e *Quantifiler™ THP DNA standard* [73].

No início de cada quantificação, preparou-se no *software* de análise dos resultados *HID Real-Time PCR Analysis Software*, a folha de planificação ou *plate layout*, servindo de orientação para a preparação das amostras a quantificar, geralmente em placa. Na sala de Pré-PCR, e na camara de fluxo destinada às amostras referência, foram preparadas diluições em série do *Quantifiler™ THP DNA Standard*, utilizadas para obtenção de uma curva de calibração, a partir da qual é estimada a concentração de DNA das amostras. Os cinco *standards*, com concentrações entre os 50 e 0,005 ng/µL, foram definidos a partir de uma solução de DNA com concentração conhecida (100 ng/µL). Estas diluições foram preparadas com o tampão *Quantifiler™ THP DNA Dilution Buffer*, em concordância com o descrito na tabela 2.1.

Tabela 2.1 Preparação das diluições em série a partir da solução *Quantifiler™ THP DNA Standard* de 100 ng/µL

Standard (Std.)	Concentração (ng/ µL)	Preparação	
		DNA	Tampão

<b>Std. 1</b>	50	10 $\mu$ L <i>DNA Standard</i>	10 $\mu$ L
<b>Std. 2</b>	5	10 $\mu$ L Std. 1	90 $\mu$ L
<b>Std. 3</b>	0,5	10 $\mu$ L Std. 2	90 $\mu$ L
<b>Std. 4</b>	0,05	10 $\mu$ L Std. 3	90 $\mu$ L
<b>Std. 5</b>	0,005	10 $\mu$ L Std. 4	90 $\mu$ L

De seguida, preparou-se a mistura de reação, adicionando-se por amostra, o volume de reagente nas quantidades que se apresentam na Tabela 2.2. A mistura de reação, é preparada para um número de amostras superior ao estudado, por forma a prevenir perdas de volume durante a pipetagem.

Tabela 2.2 Componentes e respetivos volumes por reação de quantificação

Componentes	Volume por reação
<i>Quantifiler® THP PCR Reaction Mix</i>	10 $\mu$ L
<i>Quantifiler® Trio Primer Mix</i>	8 $\mu$ L
DNA	2 $\mu$ L

Seguidamente, dispensou-se 18  $\mu$ L da mistura de reação em cada poço da placa, depois de agitar a mistura no vórtex, e adicionados 2  $\mu$ L de cada amostra ou *standard* no poço respetivo, segundo o esquema da placa previamente elaborado, perfazendo um volume de 20  $\mu$ L. Faz parte do procedimento interno da metodologia de quantificação da Delegação do Sul do SGBF, serem aditados três controlos positivos com diferentes concentrações: 0,1, 10 e 100 ng/ $\mu$ L.

Finda a preparação da placa, a mesma foi selada com uma película transparente, valendo-se de um utensílio apropriado, para evitar a formação de rugosidades que possam afetar a deteção da fluorescência. Posteriormente, centrifugou-se a placa, de forma a remover quaisquer bolhas que se tenham formado e que interferem com a amplificação e, de seguida, introduzimo-la no equipamento *7500 Real-Time PCR*. O programa de amplificação conducente à quantificação do DNA na amostra é: 95°C durante 2 minutos, seguida de 40 ciclos de 95°C durante 9 segundos e 60°C por 30 segundos. Uma vez concluída a quantificação, os resultados obtidos são analisados por meio do *software* de análise *HID Real-Time PCR Analysis Software v1.2 (Thermo Fisher Scientific)*, que permite estimar a concentração DNA em cada amostra através de análise de regressão, usando as amostras padrão como referência, que gera as curvas de calibração e calcula as quantidades de DNA em cada amostra.

## 2.6 Amplificação por PCR *multiplex*

A reação em cadeia de polimerase *multiplex*, permite a amplificação simultânea de diversos fragmentos de DNA, numa única reação de PCR, com recurso a vários pares de *primers* marcados com diferentes fluorocromos. Desta forma, é-nos possível obter informação adicional com um único teste, ao mesmo tempo reduzindo o tempo e os recursos gastos.

De uma forma geral, as reações de amplificação dividem-se em três fases, a desnaturação do DNA, a ligação dos *primers* e, finalmente, a extensão das novas cadeias pela adição de dNTPs presentes na mistura de reação, por ação da DNA polimerase. As condições da reação, temperatura, tempo e número de ciclos, varia consoante o tipo de marcadores estudados e consoante as particularidades dos reagentes utilizados, que quando aplicados em *Kit*, as condições são indicadas pelo respetivo fabricante.

O *Kit Investigator® DIPplex (QIAGEN)* foi neste trabalho aplicado para o estudo dos marcadores InDel. Para o estudo de marcadores do tipo STR, utilizou-se o *Kit PowerPlex® ESI 17 (Promega)*, recorrendo aos *Kits PowerPlex® Fusion 6C (Promega)* e *Globalfiler® (Thermo Fisher)*, para confirmação de alelos microvariante raros.

### 2.6.1 Estudo de marcadores do tipo InDel

O *Kit Investigator® DIPplex* fornece todos os reagentes necessários à amplificação dos 30 marcadores InDel e da Amelogenina, marcados com quatro fluorocromos diferentes – 6-FAM, BTG, BTY e BTR –, que se distribuem por 19 cromossomas e a pelo menos 10Mpb de distância de qualquer marcador STR e SNP comercialmente disponível.

Para o ensaio de amplificação *multiplex*, foi preparada uma *master mix*, com os reagentes e volumes indicados na tabela 2.3.

Tabela 2.3 Componentes e respetivos volumes de preparação da *master mix* para amplificação com o *Kit Investigator® DIPplex (QIAGEN)*

Componentes	Volume por reação
H <sub>2</sub> O desionizada	6,2 µL
<i>Reaction Mix A</i>	2,5 µL
<i>Primer Mix</i>	2,5 µL
<i>Multi Taq2 DNA Polymerase</i>	0,3 µL
DNA	1 µL

A mistura de reação, uma vez preparada, foi agitada por alguns segundos no vórtex, tendo-se de seguida distribuído 11,5 µL da *master mix*, pelos respetivos tubos. A quantidade de DNA adicionada aos tubos, variou entre 1 e 7,2 µL, consoante os resultados da quantificação, ajustando-se o volume de água para um volume total de reação de 12,5 µL. Relativamente aos controlos, positivo e negativo, os mesmos foram realizados adicionando 1 µL de DNA controlo 9947 e água livre de nucleases, respetivamente. Posteriormente, as amostras foram brevemente centrifugadas e colocadas no termociclador *Geneamp™ 9700 PCR System (Applied Biosystems™)* para amplificar, de acordo com as indicações do fabricante, apresentadas na tabela 2.4.

Tabela 2.4 Protocolo de amplificação do *Kit Investigator® DIPplex (QIAGEN)*

Temperatura	Tempo	Número de ciclos
94°C	4 minutos	-
94°C	30 segundos	28 ciclos
61°C	120 segundos	
72°C	75 segundos	
68°C	60 minutos	-
10°C	∞	-

### 2.6.2 Estudo de marcadores do tipo STR

A preparação da mistura de reação para amplificação dos marcadores STR, sucedeu-se de forma semelhante ao descrito na seção anterior para o estudo dos marcadores InDel, variando apenas, os componentes e respetivos volumes de preparação da *master mix* e os protocolos de amplificação. A informação relativa a cada um dos *kits* utilizados, *PowerPlex® ESI 17 (Promega)*, *PowerPlex® Fusion 6C (Promega)* e *GlobalFiler® (Thermo Fisher)*, apresenta-se nas tabelas 2.5, 2.6 e 2.7, respetivamente.

Todas as amostras neste estudo foram tipadas com o *kit PowerPlex® ESI 17 (Promega)*, que permite a co-amplificação de todos os marcadores STR do Painel Europeu, e dos *loci* D19S433, SE33, D16S539 e D2S1338. Os alelos que não estão incorporados no marcador alélico do referido *kit*, sendo detetados como picos *off-ladder*, foram confirmados pela reamplificação dessas amostras recorrendo aos *kits PowerPlex® Fusion 6C (Promega)* – *locus* D2S441 e *GlobalFiler® (Thermo Fisher)* – *loci* D2S441 e SE33. Após a reação de amplificação, o produto amplificado pode ser guardado a 4°C, durante

um período de tempo reduzido, a -20°C durante um intervalo de tempo mais alargado ou seguir logo para eletroforese capilar.

Tabela 2.5 Componentes e respetivos volumes de preparação da *master mix* para amplificação com o kit *PowerPlex® ESI 17 (Promega)*

Componentes do kit	Volume por reação	Protocolo de amplificação		
H <sub>2</sub> O desionizada	7,75 µL	96°C	2 minutos	-
<i>Reaction Mix</i>	2,5 µL	94°C	30 segundos	30 ciclos
<i>Primer Mix</i>	1,25 µL	59°C	2 minutos	
DNA alvo/controlo 9947A	1 µL	72°C	90 segundos	
<b>Volume total</b>	12,5 µL	60°C	45 minutos	-
-	-	4°C	∞	-

Tabela 2.6 Componentes e respetivos volumes de preparação da *master mix* para amplificação com o kit *PowerPlex® Fusion 6C (Promega)*

Componentes do kit	Volume por reação	Protocolo de amplificação		
H <sub>2</sub> O desionizada	6,5 µL	96°C	1 minutos	-
<i>Reaction Mix</i>	2,5 µL	96°C	5 segundos	28 ciclos
<i>Primer Mix</i>	2,5 µL	60°C	1 minutos	
DNA alvo/controlo 2800M	1 µL	60°C	10 segundos	
<b>Volume total</b>	12,5 µL	4°C	∞	-

Tabela 2.7 Componentes e respetivos volumes de preparação da *master mix* para amplificação com o kit *Globalfiler® (ThermoFisher)*

Componentes do kit	Volume por reação	Protocolo de amplificação		
<i>Low TE Buffer</i>	6,5 µL	95°C	1 minutos	-
<i>Master Mix</i>	3,75 µL	95°C	1 minutos	-
<i>Primer Mix</i>	1,25 µL	94°C	10 segundos	29 ciclos
DNA alvo/controlo 007	1 µL	59°C	90 segundos	-
<b>Volume total</b>	12,5 µL	60°C	10 minutos	-
-	-	4°C	∞	-

## 2.7 Eletroforese Capilar do produto amplificado

A separação e deteção dos diversos fragmentos de DNA, presentes no produto amplificado obtido, ocorreu por eletroforese capilar, com recurso ao sequenciador automático *Genetic Analyzer 3130 xl (Applied Biosystems)*. Este método, inicialmente proposto por Jorgenson e Lukacs, em 1981, permite a separação dos fragmentos de DNA de acordo com o seu tamanho e deteção com marcação por fluorescência.

O DNA é uma molécula negativamente carregada, em condições de pH neutro, devido ao grupo fosfato que possui. Durante a eletroforese capilar, os fragmentos de DNA migram do polo negativo (ou seja, do cátodo) para o polo positivo (o ânodo), por influência de um campo elétrico. Os referidos eletrodos, cada um no seu respetivo recipiente, que contém o tampão aquoso, encontram-se ligados a

uma fonte de energia de alta voltagem, ajudando a induzir a formação da corrente elétrica, responsável pela movimentação dos fragmentos de DNA, separados consoante o tamanho. Os fragmentos de tamanho mais reduzido, migram a uma velocidade superior aos fragmentos de tamanho maior, que ficam mais facilmente retidos nos poros do polímero, dificultando e atrasando o seu deslocamento [59,74].

Antes de proceder à eletroforese capilar propriamente dita, tanto no estudo dos marcadores InDel, quanto para os marcadores do tipo STR, as amostras foram preparadas numa placa de 96 poços, *MicroAmp® Optical* (*Applied Biosystems*). Para ambos os tipos de marcadores, começou-se por preparar a mistura de reação, contendo formamida Hi-Di™ (*Applied Biosystems*) e o marcador interno de peso molecular, que para análise dos produtos de amplificação dos InDels é o 550 (BTO) (*QIAGEN*), e para os produtos de amplificação dos STRs, o CC5 500 (*Promega*) – kits *PowerPlex® ESI 17* (*Promega*) e *PowerPlex® Fusion 6C* (*Promega*) – e o GeneScan™ 600 Liz® (*ThermoFisher*) – kit *Globalfiler®* (*ThermoFisher*). A formamida tem como função, quebrar as pontes de hidrogénio entre as cadeias complementares dos fragmentos de DNA. Na tabela 2.8, apresentam-se os volumes adicionados de cada reagente, por amostra estudada.

Tabela 2.8 Reagentes necessários para a preparação do produto amplificado para eletroforese capilar e respetivos volumes, por cada reação.

Marcadores do tipo InDel		Marcadores do tipo STR ( <i>Promega/ThermoFisher</i> )	
Reagentes	Volume por reação	Reagentes	Volume por reação
Formamida Hi-Di™	11,4 µL	Formamida Hi-Di™	10 µL / 9,6 µL
<i>Size standard</i> 550 (BTO)	0,4 µL	<i>Size standard</i> CC5 500/ GeneScan™ 600 Liz®	0,5 µL / 0,4 µL
<b>Volume Total</b>	11,9 µL	<b>Volume Total</b>	10,5 µL / 10 µL

A cada poço foi adicionado 11,9 µL, 10,5 µL ou 10 µL da mistura de reação, depois de agitada por breves segundos no vortex, e 1 µL de produto amplificado ou de *ladder* alélica. O marcador interno de peso molecular, é constituído por fragmentos de DNA de tamanhos conhecidos, providenciando os pontos de referência para determinar o tamanho dos fragmentos de DNA de uma amostra. Por sua vez, o *ladder* alélico, possui os alelos mais comuns na população em geral, já pré-amplificados, permitindo a identificação dos fragmentos de DNA desconhecidos aquando da análise. Após desnaturação durante cerca de 3 minutos a 95°C no termociclador *GeneAmp PCR system 9700* (*Applied Biosystems*), realizou-se a eletroforese capilar, nas condições especificadas pelos fabricantes dos *kits*, utilizando o polímero POP4® (*Applied Biosystems*).

## 2.8 Detecção e análise do produto amplificado

Os *primers* utilizados no processo de amplificação dos fragmentos de DNA, tanto no estudo dos marcadores InDel como nos STR, encontram-se marcados com diferentes fluorocromos, que absorvem a comprimentos de onda semelhantes, mas emitem fluorescência a diferentes comprimentos de onda. Isto permite que um único laser seja capaz de excitar quatro ou mais fluorocromos, aquando da passagem dos fragmentos de DNA pelo feixe de luz.

A emissão da fluorescência é quantificada e detetada graças a um detetor acoplado ao capilar, tratando-se de uma deteção multicolor, em que todas as cores e, como tal, todos os *loci*, podem ser detetados e distinguidos numa única injeção capilar, visto que todos os fluorocromos emitem fluorescência a diferentes comprimentos de onda [75]. O sinal de fluorescência é convertido em dados digitais, obtidos sob a forma de eletroforegramas, por sua vez armazenados num formato compatível com o software de análise, tendo-se para isso recorrido ao software GeneMapper® ID-X versão 1.4. (*Applied Biosystems*) [75].

## 2.9 Análise estatística

Para análise estatística dos genótipos obtidos no estudo dos marcadores InDel e STR, foram utilizados diferentes programas informáticos, consoante os parâmetros de interesse. No que diz respeito ao cálculo dos parâmetros, heterozigotia observada e esperada, frequências alélicas e HWE, utilizou-se o *software* Arlequin versão 3.5.1.2, amplamente utilizado para estudos de diversidade genética, recorrendo a algoritmos de simulação, o que permite estimar os parâmetros genéticos populacionais de interesse [76]. Em relação aos parâmetros de interesse forense, para os marcadores do tipo InDel, calculou-se o poder de discriminação e exclusão e para além destes, foram também calculados para os marcadores STR, a probabilidade de *matching*, PIC e índice típico de paternidade, recorrendo ao *software* PowerStats versão 12.

## 2.10 Análise Filogenética

A população Brasileira residente em Lisboa, foi no estudo dos marcadores InDel, comparada com as populações portuguesa, a população imigrante de Cabo-Verde, Angola, Moçambique, Guiné-Bissau e de São Tomé e Príncipe, anteriormente caracterizadas para esse mesmo tipo de marcadores genéticos. A matriz de distâncias foi calculada com recurso ao *software* Arlequin versão 3.5.1.2, sendo depois utilizada para obter a representação gráfica das distâncias genéticas entre as populações estudadas, utilizando o *software* Mega7, através do método *neighbor joining* (NJ) [76,77].

# Capítulo 3 – Resultados e Discussão

## 3.1 Estudo dos Marcadores InDel

A análise da figura 3.1, revela um electroforegrama de uma amostra biológica analisada com o *kit Investigator® DIPplex (QIAGEN)*. Nesta amostra é possível observar, a deteção dos 30 InDel e do gene da Amelogenina, em quatro painéis diferenciados cromaticamente, de acordo com a marcação fluorescente dos *primers* fornecidos pelo *kit*, isto é, a cor de cada painel – verde, azul, amarelo (que na imagem surge a preto para facilitar a sua interpretação) e vermelho -, corresponde ao comprimento de onde a que o fluorocromo emite fluorescência.



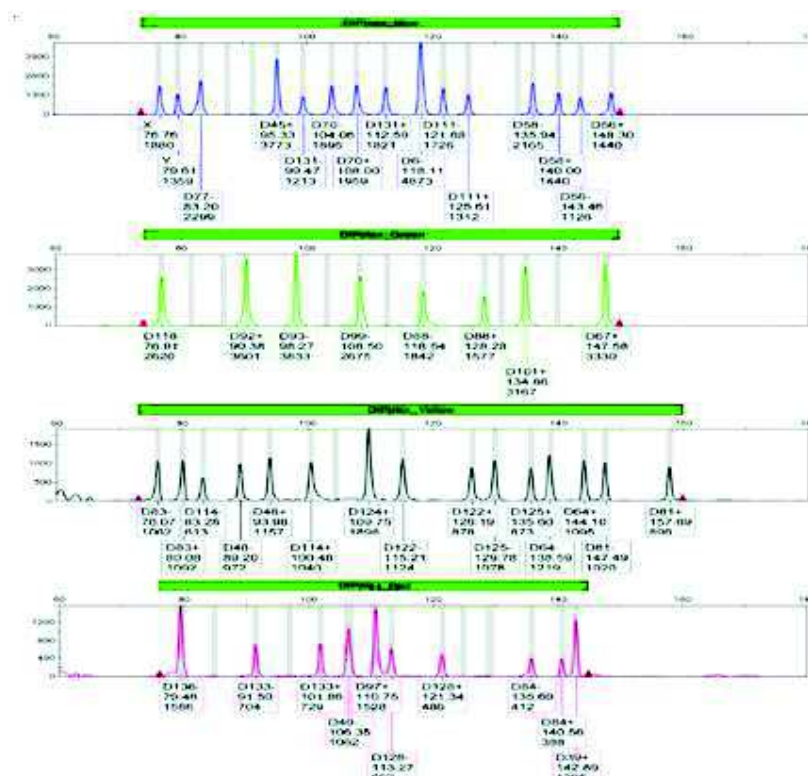


Figura 3.1 Exemplo de um eletroforegrama obtido aquando da detecção e análise dos marcadores do tipo InDel: Genótipo de um indivíduo do sexo masculino oriundo do Brasil. Cada painel contém 7 a 8 marcadores, correspondendo a um máximo de 14 a 16 picos.

Para todos os 244 indivíduos estudados com estes marcadores, foi possível obter um perfil completo, registrando-se, contudo, a existência de um perfil com desbalanceamento no tamanho do pico no *locus* HLD97. Apesar de não se encontrarem previstos pelo fabricante, certos autores têm descrito a ocorrência de desbalanceamentos e *drop-out* na utilização do *Kit Investigator® DIPplex (QIAGEN)*.

Em 2011, Fondevila e colaboradores, observaram desbalanceamento entre picos heterozigóticos nos *loci* HLD97 e HLD83, sugerindo que a dificuldade na amplificação resultaria da presença de polimorfismos na região de ligação dos *primers*, porém, a sequenciação do *locus* HLD83 não revelou a existência de tais variações. Quanto ao *locus* HLD97, este será discutido mais ao pormenor. Foram ainda detetadas pelos mesmos autores, microvariantes móveis nos *loci* HLD99 e HLD84, tendo os autores apontado como explicação mais provável, a existência de um segundo InDel que poderia afetar o tamanho do fragmento amplificado, desviando a mobilidade do pico na eletroforese. Estas variações foram igualmente descritas por LaRue *et al.* (2012), também aquando do estudo de quatro grupos humanos representativos da população Estadunidense [78,79].

Friis *et al.* (2012), identificaram na população dinamarquesa, uma microvariante no *locus* HLD131, resultante de um polimorfismo na região de ligação dos *primers*. Kis e colaboradores (2012), reportaram na população Húngara, desbalanceamentos nos *loci* HLD6 e HLD136, enquanto, Zidkova e colaboradores (2013), o fizeram para os *loci* HLD6, HLD40 e HLD58, ao estudarem a população Checa. No ano seguinte, em 2014, num estudo populacional da Coreia do Sul, Seong *et al.* verificaram a existência de uma microvariante no *locus* HLD93. Também no *locus* HLD128, Tomas e colaboradores, em 2015, reportaram a existência de um desbalanceamento na população Grega [80–82].

Apesar dos fenómenos terem sido detetados em estudos da população Europeia, bem como em estudos acerca das populações dos Estados Unidos e Coreia do Sul, não nos é possível descartar a possibilidade da ocorrência de fenómenos semelhantes, nestes e noutros *loci*, em indivíduos da



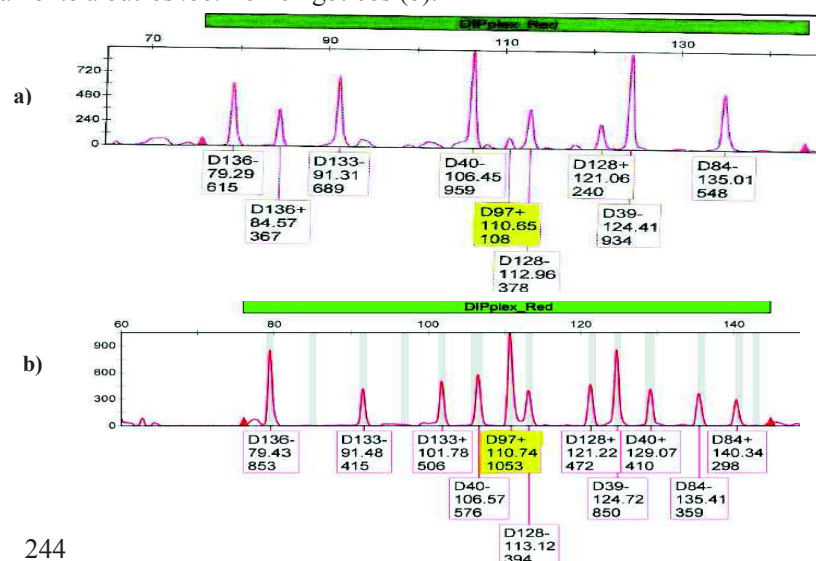
população Brasileira. Particularmente em genótipos de indivíduos cuja taxa de amplificação, de um modo geral, foi menor, dificultando a identificação de *drop-out* devido ao tamanho reduzido dos picos.

### 3.1.1 Microvariantes

A partir da análise dos eletroforegramas obtidos, foi-nos possível assinalar dois tipos de variações alélicas que influenciam a interpretação dos genótipos, uma no *locus* HLD97 a outras nos *loci* HLD99 e HLD56, nenhuma descrita pelo fabricante do *Kit Investigator® DIPplex*.

#### 3.1.1.1 – *Locus* HLD97

No *locus* supramencionado, em homozigotia, observou-se um desbalanceamento do tamanho do pico correspondente ao alelo de inserção (D97+), muito diminuído face a outros sistemas homozigóticos. A Figura 3.2 permite-nos observar a referida microvariante (a), bem como o exemplo de um indivíduo onde o alelo foi detetado, também em homozigotia, sem alterações e de forma equilibrada comparativamente a outros *loci* homozigóticos (b).



Dos 244 indivíduos com  
 Figura 3.2 Diferenças no tamanho (RFUs) dos picos do alelo correspondente à inserção do marcador HLD97 (D97+): a) pode-se verificar a presença de um pico com 108 RFUs no local correspondente ao referido alelo e ausência de pico para o alelo de deleção D97-, indicando tratar-se de um indivíduo homozigótico para o *locus* em questão, apesar do tamanho reduzido do pico relativamente aos restantes; b) exemplo de um indivíduo homozigótico D97+, onde o alelo foi detetado de forma balanceada.

nacionalidade e ancestralidade Brasileira estudados, apenas um apresentou esta variação alélica, correspondendo a uma frequência de 0,004. A repetição da genotipagem do indivíduo em questão, permitiu verificar a reprodutibilidade desta microvariante, anteriormente reportada noutros estudos populacionais [27,78,80].

Nessas populações, Fondevila e colaboradores (2011, 2012), Friis e colaboradores (2012) e Martín e colaboradores (2013), para confirmação do genótipo dos indivíduos e estudo da microvariante, procederam à sequenciação da região onde a mesma se localiza. A suspeição dos autores, acerca da existência de um polimorfismo na região de ligação dos *primers*, que impedia a correta amplificação e consequente diminuição no tamanho do pico para o alelo D97+, foi dessa forma confirmada, permitindo verificar em todos os casos a presença de um SNP (A/G). O polimorfismo trata-se da substituição de uma guanina por uma adenina, 61pb a jusante do *locus* HLD97, entre os 97pb e 110pb do fragmento amplificado. Os autores observaram ainda, que os fragmentos portadores do Alelo A, apresentavam uma

taxa de amplificação diminuída relativamente aos portadores da variante G, diferença refletida na redução e desbalanceamento do tamanho do pico detetado para o alelo de inserção D97+.

Embora previamente reportada em Afro-americanos [78,79], Dinamarqueses [80], Iranianos [83], Espanhóis [27] e em populações nativas Bantu – Xhosa e Zulu – [84], para nosso conhecimento, a microvariante resultante do polimorfismo rs17245568, não foi anteriormente descrita em estudos da população Brasileira, que se caracteriza como uma população altamente miscigenada.

Em dois estudos de avaliação do *Kit Investigator® DIPplex* com quatro populações dos Estados Unidos – afro-americana, caucasiana, hispânica e asiática –, foi na primeira que Fondevila e colaboradores e LaRue e colaboradores, observaram com maior frequência a variante de altura de pico reduzida para o alelo D97+ (0,22 e 0,28, respetivamente). Por outro lado, em ambos os estudos, foi na população caucasiana que a microvariante foi menos frequentemente observada (0,04 e 0,044, respetivamente) [78,79]. No estudo da população Iraniana com 253 indivíduos, quatro apresentavam fraca deteção de sinal para o alelo D97+, em homozigotia, correspondendo a aproximadamente 1,6 % [83]. No estudo com populações nativas Bantu, a frequência do alelo nulo D97+ foi de 0,26 [84], já nos estudos realizados na população Espanhola e Dinamarquesa, não foram apresentadas frequências para o alelo microvariante.

### 3.1.1.2 – *Loci* HLD99 e HLD56

Nos *loci* acima referidos, foram identificadas microvariantes móveis, pela deteção de picos *off ladder* que apresentavam a mesma força de sinal que os picos dos restantes *loci* analisados pelo *Kit*.

Em relação ao *locus* HLD99, em heterozigotia, verificou-se que o pico correspondente ao alelo de inserção apresentava um nucleótido a menos do que o descrito pelos fabricantes do *Kit*, isto é, aproximadamente 112 pb ao invés de 113 pb, tendo sido designada a microvariante por D99+(-1) (ponto a), figura 3.3). Esta microvariante foi detetada em dois dos indivíduos estudados. Relativamente ao *locus* HLD56, constatou-se que o pico para o alelo de deleção apresentava dois nucleótidos a mais, ou seja, cerca de 145 pb em vez de 143 pb, neste caso designando-se a microvariante por D56-(-2) (ponto c), figura 3.3). Ambas as microvariantes, D99+(-1) e D56-(-2) ocorreram, respetivamente, com uma frequência de 0,004 e 0,002 na população Brasileira residente em Portugal, à qual dedicamos o nosso estudo.

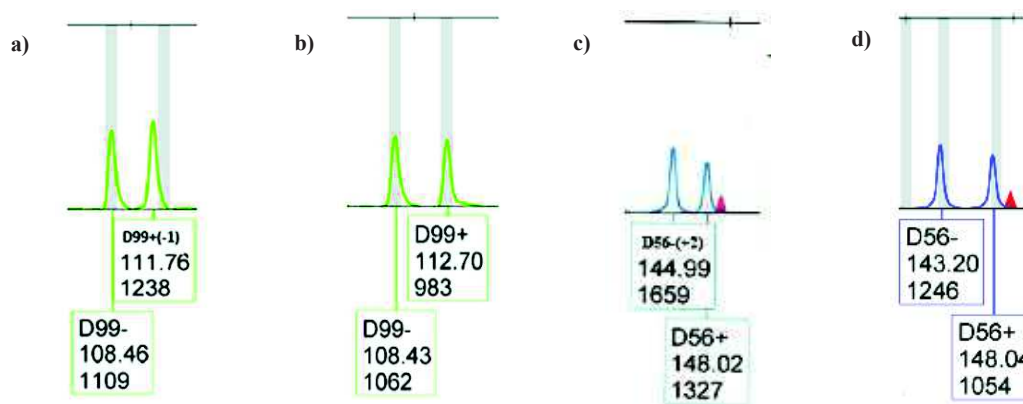


Figura 3.3 Microvariantes detetadas quando da análise dos marcadores InDel: a) pico *off ladder* assinalado como D99+(-1), detetado no painel correspondente à cor verde, com um nucleótido a menos do que b) o alelo D99+ descrito; c) pico *off ladder* assinalado como D56-(-2), detetado no painel correspondente à cor azul, com dois nucleótidos a mais que d) o alelo D56- descritos.

A repetição da tipagem dos indivíduos em quem foram detetadas as microvariantes, permitiu confirmar o genótipo dos *loci*, uma vez verificada a reprodutibilidade dos alelos. Contudo, tal como

referido na secção anterior, uma forma de estudar e caracterizar ambas as microvariantes, e também a melhor, seria através da sequenciação das regiões onde se apresentam estes alelos.

A variação alélica detetada no *locus* HLD99, já se encontra bem descrita e estudada por alguns autores, que através da sequenciação da região, provaram a existência de deleção de uma adenina, 4 pb a montante do alelo para a inserção. É a este polimorfismo, referenciado no banco de dados público dbSNP como rs11346981, que se deve a alteração da mobilidade e, consequentemente, o aparecimento de um pico *out off ladder*, que apesar de 1 pb mais pequeno, corresponde ao alelo de inserção do *locus* HLD99 [78,79]. Localizada na região em estudo, esta microvariante apresenta-se como um terceiro alelo de tamanho diferente do *locus* em questão, sendo também ela considerada um polimorfismo InDel. Esta caracterização de variantes móveis, dota o *Kit Investigator® DIPplex*, de um maior poder de discriminação entre indivíduos [78,79].

Em relação à microvariante detetada no *locus* HLD56, até o presente momento, nenhum outro estudo realizado nesta ou noutras populações, relatou a observação desta variação alélica [62]. Por conseguinte, de modo a caracterizá-la da melhor forma, dever-se-ia proceder à sequenciação do polimorfismo em questão.

### 3.2 Estudo dos marcadores STR

Todas as amostras no presente estudo foram tipadas com recurso ao *Kit PowerPlex® ESI 17 (Promega)*, que permite a co-amplificação de 17 *loci*, cuja deteção e análise, analogamente ao estudo dos marcadores InDel, foi elaborada por eletroforese capilar com deteção por fluorescência. Os resultados obtidos, apresentam-se sobre a forma de um eletroforegrama, em que os picos correspondem aos alelos STR amplificados. Os genótipos dos indivíduos, foram desta forma obtidos, tendo-se considerado apenas alelos com sinal, em RFUs – do inglês, *Relative Fluorescence Units* –, a partir de 300.

Dos 244 indivíduos de origem Brasileira estudados, apenas um apresentava um perfil genético parcial, excluído da análise estatístico dos dados. Adicionalmente, verificou-se em dois dos indivíduos, a presença de um alelo *off-ladder*, um no *locus* SE33 e o outro no *locus* D2S441, que correspondem a microvariantes alélicas que não se encontram previstas no *kit PowerPlex® ESI 17 (Promega)*. Estas variações alélicas, serão discutidas em mais pormenor, na secção que se segue.

#### 3.2.1 Microvariantes *off-ladder*

No presente estudo foram detetados dois picos *off-ladder*, em indivíduos diferentes, com força de sinal idêntica aos picos dos restantes *loci* analisados pelo *kit*, e que correspondem a alelos microvariante, que não se encontram previstos no *ladder* alélico. Estas variantes alélicas, tratam-se do alelo 21.2 para o *locus* SE33 e 14.1 para o *locus* D2S441, que ocorreram com uma frequência de 0,005 e 0,002, respetivamente.

O marcador SE33 destaca-se como um *locus* STR hipervariável, presente no braço longo do cromossoma 16, na banda 4 da região 1, que contém uma estrutura de repetição complexa AAAG, que se repete 3 a 49 vezes [59,85]. O alelo microvariante 21.2, detetado pelo *kit PowerPlex® ESI 17 (Promega)* como um pico *off-ladder*, foi confirmado pela repetição da tipagem do indivíduo em questão com o *kit Globalfiler® (Thermo Fisher)*, comprovando-se a reprodutibilidade do alelo, previsto no marcador alélico do *kit* (Pontos a) e b), Figura 3.4). Rodenbusch e colaboradores (2009), num estudo da população Brasileira, encontraram esta variante alélica com uma frequência de 0.0078, relativamente semelhante à verificada no presente trabalho [86].

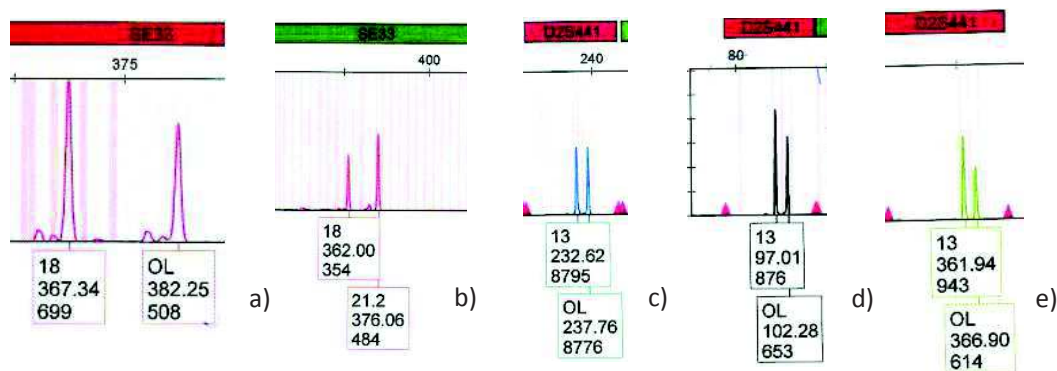


Figura 3.4 **Microvariantes detetadas aquando da análise dos marcadores STR:** a) pico *off ladder* detetado no marcador SE33, utilizando o kit *PowerPlex® ESI 17*, confirmando-se, pela repetição da tipagem do indivíduo em questão com o kit *GlobalFiler®* b), que se trata do alelo microvariante 21.2; c) pico *off ladder* detetado no marcador D2S441 com o kit *PowerPlex® ESI 17*, comprovando-se a reprodutibilidade do alelo microvariante raro 14.1, pela repetição da tipagem do indivíduo, recorrendo aos kits c) *GlobalFiler®* e d) *PowerPlex® Fusion 6C*.

O locus D2S441 presente na banda 4 da região 1, no braço curto do cromossoma 2, é um STR tetranucleotídico com sequência de repetição TCTA, que se repete 8 a 17 vezes [59,87]. O alelo microvariante raro 14.1, detetado como um pico *off-ladder* para o referido marcador, foi confirmado pela repetição da tipagem do indivíduo a quem foi identificado, recorrendo aos kits *PowerPlex® Fusion 6C* (Promega) e *Globalfiler®* (Thermo Fisher). Embora a variante alélica 14.1, não se encontre prevista nos *ladders* alélicos dos kits acima mencionados, os resultados obtidos permitiram confirmar a sua reprodutibilidade (Pontos c), d) e e), Figura 3.4).

Num estudo com 2085 indivíduos da população Holandesa, A. Westen e colaboradores (2014), identificaram esta variante alélica num dos genótipos obtidos, o que corresponde a uma frequência de 0,00024. O alelo que poderá resultar da deleção de 3 pb na unidade de repetição, também já foi reportado num estudo anterior da população Brasileira com cerca de 96 400 indivíduos, apresentando-se com uma frequência de 0,000047 [88,89].

O desempenho destes marcadores, e a capacidade dos kits detetarem variantes alélicas, constituem mais-valias para a identificação individual, bem como investigações de filiação e casos complexos de paternidade reversa.

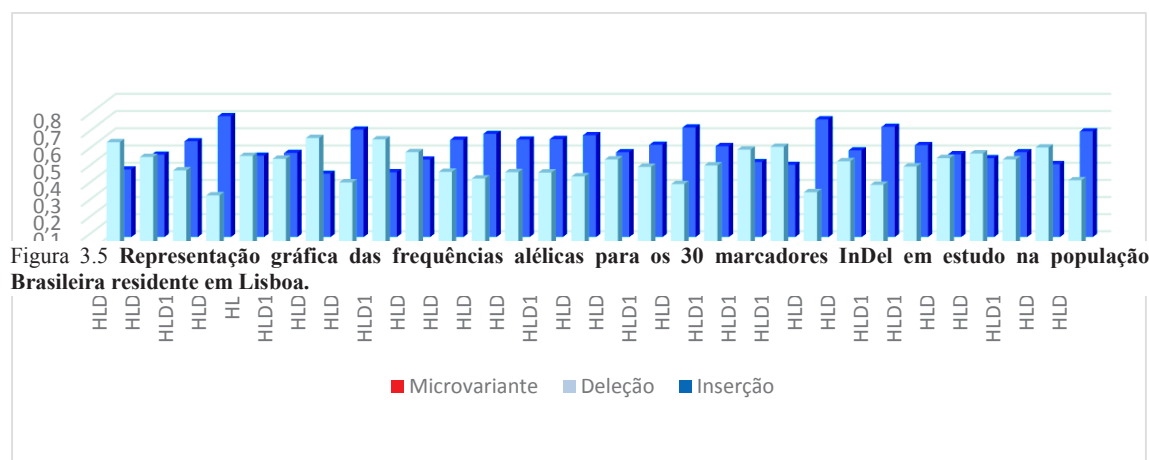
### 3.3 Estudo da população imigrante Brasileira residente em Lisboa

Para a caracterização genética da população imigrante do Brasil a residir em Lisboa, foram tipados 244 indivíduos, todavia, os *loci* STR apresentam-se em menor número comparativamente aos marcadores InDel, já que não foi possível obter perfil genético completo para um dos indivíduos estudados. Os resultados obtidos relativamente aos diversos parâmetros calculados, encontram-se descritos na tabela 3.1.

As frequências alélicas calculadas para os 30 marcadores InDel apresentam, para certos *loci*, algum desbalanceamento entre as frequências de inserção e deleção, sendo mais evidente a dissimelhança nos *loci* HLD70 e HLD64, tal como nos é permitido observar pela figura 3.5. Em relação ao *locus* HLD70, o alelo correspondente à inserção apresenta uma frequência de 0,705 e o alelo de deleção 0,295. Já o *locus* HLD64 exhibe, para o alelo de inserção, uma frequência de 0,686 e 0,314 para o alelo de deleção. Zidkova e colaboradores, em 2013, confirmaram a ocorrência de *drop-out* nos *loci* HLD40, HLD6 e HLD58, que pode estar associado ao desbalanceamento das frequências alélicas [81]. Assim, e apesar do *drop-out* nos *loci* HLD70 e HLD64 não se encontrar ainda descrito na literatura, a sua existência é uma possibilidade a considerar, tendo em conta o desequilíbrio verificado entre as frequências de inserção e deleção. Num estudo da população Brasileira, Palha e colaboradores, em 2015, verificaram que os *locus* HLD64, HLD70 e HLD48, eram os que apresentavam maior desbalanceamento entre frequências alélicas [65].

A figura 3.5 permite-nos ainda observar, que os *loci* InDel com menor desbalanceamento entre frequências alélicas, correspondem ao HLD128 e HLD83, cujos alelos de inserção e deleção, apresentam uma frequência de 0,496 e 0,504, respetivamente. Os *loci* HLD111, HLD81, HLD45 e HLD114, exibem igualmente frequências alélicas próximas de 0,500.

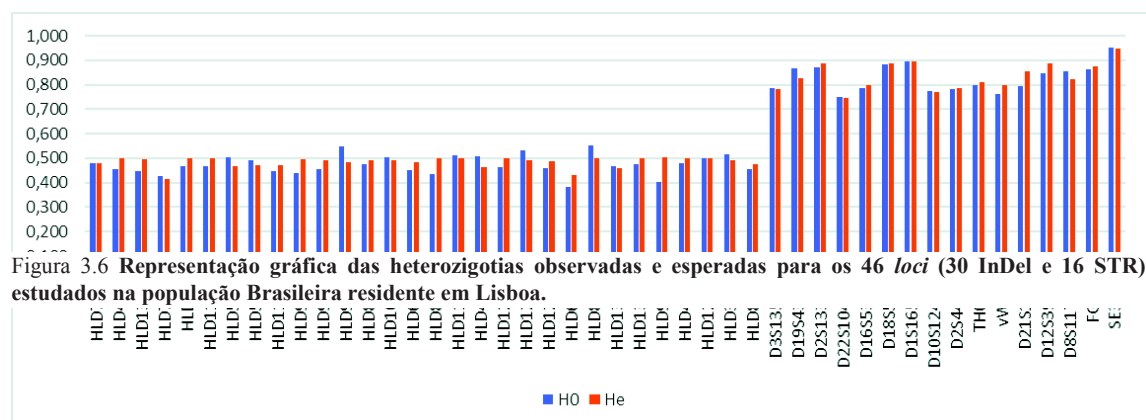
Neste estudo, verificou-se a existência de microvariantes nos *loci* HLD97, HLD99 e HLD56, consideradas como um terceiro alelo, aumentando o poder de discriminação do painel de marcadores.



As microvariantes no *locus* HLD97 e HLD99, como anteriormente referido, encontram-se descritos na literatura, mas não para a população Brasileira [79,90]. Relativamente ao *locus* HLD97, verificou-se na forma homozigótica, um desbalanceamento do pico correspondente ao alelo de inserção D97+, detetado num indivíduo, apresentando uma frequência de 0,004. Foram igualmente detetadas em heterozigotia, as microvariantes móveis D99+(-1) e D56-(-2), detetadas 2 e 1 vez, respetivamente, com uma frequência de 0,004 e 0,002.

Quanto aos 16 marcadores STR analisados com o *kit* PowerPlex® ESI 17 (Promega), apresenta-se na tabela 3.1, o intervalo das frequências alélicas calculadas para cada um dos *loci*. Adicionalmente, na seção anexos, figura 5.1, encontram-se as representações gráficas das frequências alélicas por marcador, cada barra dos gráficos, etiquetada com o respetivo valor da frequência. No presente estudo foi detetado a um dos indivíduos, um alelo microvariante raro 14.1 no marcador D2S441, com uma frequência de 0,002, que não se encontrava previsto no *ladder* alélico de nenhum dos *kits* aplicados para confirmação da microvariante, anteriormente reportada num estudo da população Brasileira [89].

Os valores de  $H_o$ , que diz respeito ao número de indivíduos heterozigóticos observados em relação ao número total de indivíduos estudados, e os valores de  $H_e$ , que representa a proporção de heterozigóticos esperada para uma população em equilíbrio de Hardy-Weinberg, variam mas pouco entre os 46 *loci* estudados (tabela 3.1 e gráfico 3.6).



HWE, cujo valor de  $p$  se manteve significativo, mesmo após a correção de Bonferroni ( $p < 0,002$ ). Neste sentido, podemos considerar pelo menos duas hipóteses para o desvio no HWE, a presença de alelos nulos para o *locus* em questão e o tamanho reduzido da amostra ( $2N=488$ ). Estudos realizados noutras populações, reportam desvios no HWE para o *locus* HLD97 [84,92]. Os restantes *loci* InDel apresentam-se em HWE, após correção de Bonferroni.

Tabela 3.1 **Dados relativos aos 46 loci estudados para a população imigrante oriunda do Brasil residente na região de Lisboa:** número de alelos estudados (2N), frequências alélicas, heterozigotia observada (Ho) e esperada (He), valor de  $p$  para o HWE.



HLD58	488	0,629	0,371		0,504	0,468	0,267
HLD56	488	0,371	0,627	0,002	0,492	0,470	0,683
HLD118	488	0,621	0,379		0,447	0,472	0,419
HLD92	488	0,547	0,453		0,439	0,497	0,071
HLD93	488	0,432	0,568		0,455	0,492	0,247
HLD99	488	0,393	0,602	0,004	0,549	0,483	0,014
HLD88	488	0,430	0,570		0,475	0,491	0,692
HLD101	488	0,428	0,572		0,504	0,491	0,694
HLD67	488	0,406	0,594		0,451	0,483	0,348
HLD83	488	0,504	0,496		0,434	0,501	0,040
HLD114	488	0,461	0,539		0,512	0,498	0,698
HLD48	488	0,361	0,639		0,508	0,462	0,127
HLD124	488	0,469	0,531		0,463	0,499	0,300
HLD122	488	0,561	0,439		0,533	0,493	0,243
HLD125	488	0,578	0,422		0,459	0,489	0,360
HLD64	488	0,314	0,686		0,381	0,431	0,076
HLD81	488	0,494	0,506		0,553	0,501	0,125
HLD136	488	0,357	0,643		0,467	0,460	0,890
HLD133	488	0,463	0,537		0,475	0,498	0,523
HLD97	488	0,512	0,484	0,004	0,402	0,505	0,000
HLD40	488	0,539	0,461		0,480	0,498	0,607
HLD128	488	0,504	0,496		0,500	0,501	1,000
HLD39	488	0,574	0,426		0,516	0,490	0,432
HLD84	488	0,383	0,617		0,455	0,474	0,587

***Marcadores STR***

Locus STR	2N	Frequências alélicas <i>Intervalo de variação</i>	Ho	He	<i>p-value</i>
D3S1358	486	0,002-0,290	0,786	0,782	0,854
D19S433	486	0,002-0,284	0,868	0,828	0,940
D2S1338	486	0,002-0,204	0,872	0,888	0,641
D22S1045	486	0,002-0,354	0,749	0,746	0,073
D16S539	486	0,002-0,274	0,786	0,798	0,734
D18S51	486	0,002-0,146	0,881	0,886	0,788
D1S1656	486	0,002-0,160	0,893	0,895	0,218
D10S1248	486	0,002-0,321	0,774	0,771	0,407

Locus STR	2N	Frequências alélicas <i>Intervalo de variação</i>	Ho	He	<i>p-value</i>
D2S441	486	0,002-0,282	0,782	0,786	0,082
TH01	486	0,002-0,245	0,798	0,809	0,000
vWa	486	0,002-0,284	0,761	0,799	0,807
D21S11	486	0,002-0,230	0,794	0,853	0,035

Tabela 3.1. – Continuação – Dados relativos aos 46 loci estudados para a população imigrante oriunda do Brasil residente na região de Lisboa: número de alelos estudados (2N), frequências alélicas, heterozigotia observada (Ho) e esperada (He), valor de *p* para o HWE.

D12S391	486	0,002-0,195	0,848	0,885	0,074
D8S1179	486	0,006-0,278	0,856	0,821	0,982
FGA	486	0,002-0,173	0,864	0,874	0,662
SE33	486	0,002-0,097	0,951	0,946	0,523

### 3.4 Análise de distâncias genéticas entre populações

A matriz de distâncias genéticas foi construída a partir das frequências alélicas calculadas para os 30 *loci* InDel na população em estudo, e outras seis populações anteriormente estudadas pelo nosso grupo, utilizando os mesmos marcadores – Portugal e imigrantes de Angola, Cabo-Verde, Moçambique, Guiné-Bissau e São Tomé e Príncipe. Na tabela 3.2 apresentam-se as distâncias genéticas calculadas pelos índices de fixação ( $F_{st}$ ), cujos valores podem variar entre 0 – interpretado como composição alélica idêntica – e 1 – ausência de partilha de alelos.

Pela análise da matriz, é-nos possível verificar, que a população imigrante oriunda do Brasil a residir em Lisboa, encontra-se, geneticamente mais próxima, da população imigrante de origem Moçambicana, para a qual a distância genética obtida foi a mais pequena, comparativamente às restantes populações, apenas 0,00209. Muito próxima da população imigrante Brasileira, está também a população caucasiana, com uma distância de 0,00883, seguida da população imigrante Angolana (0,02168) e Cabo-verdiana (0,02315). É em relação à população de imigrantes de São Tomé e Príncipe, com uma distância genética de 0,04608, e mais ainda da população imigrante da Guiné-Bissau, com uma distância de 0,05349, que a população em estudo se encontra mais afastada em termos genéticos.

Tabela 3.2 Matriz de distâncias genéticas entre a população em estudo - imigrantes Brasileiros residentes em Lisboa -, e entre seis populações anteriormente estudadas com os mesmos marcadores genéticos - Portugal e imigrantes de Angola, Moçambique, Guiné-Bissau, Cabo-verde e São Tomé e Príncipe.

Populações	Angola	Moçambique	Guiné-Bissau	Portugal	Cabo-Verde	São Tomé e Príncipe	Brasil
<b>Angola</b>	0,00000	*	*	*	*	*	*
<b>Moçambique</b>	0,00839	0,00000	*	*	*	*	*
<b>Guiné-Bissau</b>	0,00956	0,03355	0,00000	*	*	*	*
<b>Caucasiano</b>	0,05281	0,02029	0,09757	0,00000	*	*	*
<b>Cabo-verde</b>	0,00342	0,01103	0,00925	0,05164	0,00000	*	*
<b>São Tomé e Príncipe</b>	0,00618	0,02947	0,00707	0,08788	0,01169	0,00000	*
<b>Brasil</b>	0,02168	0,00209	0,05349	0,00883	0,02315	0,04608	0,00000

A proximidade da população Brasileira com a população Portuguesa – caucasiana – e com a população Moçambicana – africana –, reflete a elevada miscigenação da população Brasileira, fruto, como anteriormente referido, das sucessivas ondas migratórias que lhe deram origem. A observação da árvore representativa das distâncias genéticas, figura 3.7, facilita a comparação e interpretação das diferenças/semelhanças entre as populações.



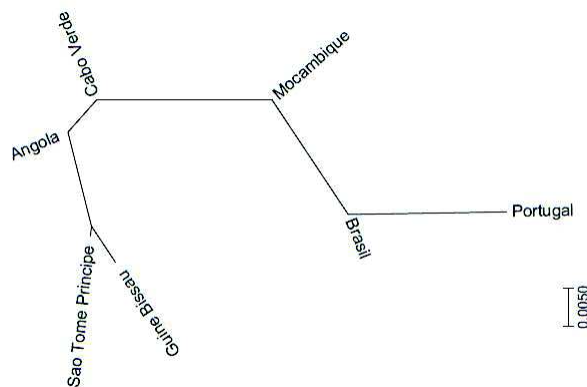


Figura 3.7 Árvore representativa das distâncias genéticas calculadas pelo índice de fixação.

A observação da figura 3.7, permite-nos confirmar a presença de quatro *clusters* distintos, um para a população Portuguesa, outro para a população imigrante Brasileira em estudo e ainda, um para a população imigrante oriunda de Moçambique, que se encontra num *cluster* diferente das restantes populações de imigrantes africanos - Angola, Cabo-Verde, Guiné-Bissau e São Tomé e Príncipe. Resumidamente, a existência de dois *clusters* para as populações africanas, deve-se, muito provavelmente, aos fenómenos migratórios do povo Bantu, um em direção a Oeste, atual região de Angola e Guiné-Bissau, outro em direção a Este, hoje território Moçambicano [93].

O arquipélago de Cabo Verde desde o seu descobrimento, por volta de 1460, foi um importante entreposto de escravos capturados da costa Oeste de África, nomeadamente das regiões de Angola – de onde terão saído cerca de 5,7 milhões de escravos entre 1501 e 1866 – e Guiné-Bissau. O arquipélago de São Tomé e Príncipe, terá sido descoberto entre 1470-1472, por Pêro Escobar, João de Paiva e João de Santarém, totalmente desabitado. Tal como Cabo Verde, também São Tomé e Príncipe serviu de entreposto de escravos africanos para tráfico negreiro, durante o Império Português, entre os séculos XV a XIX, o que poderá explicar a proximidade genética destas quatro populações – Angola, Guiné-Bissau, Cabo Verde e São Tomé e Príncipe –, que se reúnem no mesmo *cluster* [94–97].

A importância de Moçambique na rota de escravos cresceu exponencialmente, após a proibição do tráfico negreiro a norte da linha do Equador, que obrigou a deslocação da busca de escravos para o Sul de Angola e costa Oriental Africana. A figura 3.7 mostra-nos claramente, a proximidade genética da população em estudo com a população imigrante oriunda de Moçambique, relativamente às restantes populações imigrantes Africanas. Uma possível explicação para este fato, é que nos séculos XVI e XVII, a fecundidade dos escravos no Brasil era baixa, visto que, devido aos maus tratos, a sua esperança de vida era também mais reduzida, e grande parte das mulheres africanas morria sem deixar descendência. Assim, apesar da maior parte dos escravos Moçambicanos ter chegado mais tarde ao Brasil – século XIX –, a expectativa de vida nessa época era maior, graças a medidas como a Lei do Ventre Livre, motivo porque terão tido mais oportunidade de contribuírem para o fundo genética da população Brasileira, relativamente a escravos de outras regiões, que chegaram ao Brasil em séculos anteriores [98,99].

Durante a colonização do Brasil, entre 1500 e 1800, cerca de 500 000 colonizadores portugueses, maioritariamente homens, contactaram com ameríndios, civilizações indígenas que ocupavam o território até à sua chegada. Adicionalmente, após essa época, cerca 6 milhões de imigrantes oficiais, provenientes de várias regiões da Europa, nomeadamente de Portugal, contribuindo para o património genético do país. Desta forma, as ligações históricas entre Portugal e Lisboa, que perduram há mais de 500 anos, poderão constituir uma explicação favorável para a proximidade genética entre a população imigrante Brasileira a residir em Lisboa e a população Portuguesa [22,23].

### 3.5 Análise de *Linkage Disequilibrium*

A aplicabilidade forense de um conjunto de marcadores, requer a análise de LD, motivo pelo qual, após a tipagem dos indivíduos em estudo, procedeu-se à referida análise. Foram realizadas 1035 comparações, entre os 46 *loci* analisados, verificando-se a existência de LD ( $p < 0,05$ ) em 112 pares de marcadores. Após a correção de Bonferroni ( $p < 0,000048$ ), verificou-se a presença de LD em seis pares de *loci*, dois dos quais entre um InDel e um STR, que se situam em cromossomas distintos – HLD88 e D1S1656, HLD40 e D22S1045. Os restantes quatro pares de marcadores em LD foram verificados entre *loci* STR, também eles presentes em cromossomas diferentes – TH01 e vWa, D18S51 e D12S391, D16S539 e FGA e D10S1248 e FGA.

Assim, e apesar dos resultados obtidos, não se considera a ocorrência de LD entre os diversos *loci*, podendo os valores de significância para este parâmetro estarem associados a desvios no HWE, para um ou ambos os *loci*. Se, por um lado, esta associação poderá explicar a presença de LD nos *loci* vWa e TH01 - com desvio no HWE -, o mesmo não poderá ser dito em relação aos restantes pares de marcadores, que se encontram em equilíbrio na população. Neste sentido, terá que haver outra explicação para este fenómeno, que poderá passar pelo tamanho da amostra em estudo, ou ainda por fenómenos de migração, seleção ou deriva genética, responsáveis por subestruturação populacional [65]. Na tabela 5.3, apresentam-se os resultados obtidos na análise de LD para os 46 *loci* em estudo.

### 3.6 Análise dos parâmetros Forenses

Para que a individualização de um indivíduo, no âmbito forense, seja realizado com maior rigor e precisão possíveis, é necessário efetuar a avaliação dos parâmetros forenses de determinado conjunto de marcadores genéticos.

Neste sentido, foram calculados parâmetros forenses para os 46 *loci* (30 InDel e 16 STR) em estudo na população imigrante Brasileira. Na tabela 3.3 apresentam-se os valores obtidos para os parâmetros probabilidade de coincidência ou probabilidade de *matching* (MP) – do inglês, *matching probability* –, poder de discriminação (PD), poder de exclusão (PE), conteúdo informativo de polimorfismo (PIC) – do inglês, *polymorphic informative content* – e índice de paternidade típico (TPI) – do inglês, *typical paternity index*. Para cada conjunto de marcadores genéticos estudado, apresentam-se a negrito e sublinhado, na tabela 3.3, os valores máximos e mínimos encontrados para os parâmetros forenses calculados.

Tabela 3.3 Parâmetros forenses calculados para os marcadores InDel (30 *loci*) e STR (16 *loci*) estudados na população imigrante brasileira a residir na região de Lisboa.

Locus	MP	PD	PE	PIC	TPI
<i>Marcadores InDel</i>					
HLD77	0,387	0,613	0,170	0,36	0,96
HLD45	0,356	0,644	0,151	0,37	0,92
HLD131	0,360	0,640	0,145	0,37	0,90
HLD70	<b><u>0,430</u></b>	<b><u>0,570</u></b>	0,131	<b><u>0,33</u></b>	0,87
HLD6	0,361	0,639	0,160	0,37	0,94
HLD111	0,360	0,640	0,160	0,37	0,94
HLD58	0,410	0,590	0,191	0,36	1,01
HLD56	0,399	0,601	0,180	0,36	0,98
HLD118	0,382	0,618	0,145	0,36	0,90
HLD92	0,354	0,646	0,139	0,37	0,89
HLD93	0,365	0,635	0,151	0,37	0,92
HLD99	0,417	0,583	0,234	0,37	1,11
Tabela 3.3 – Continuação – Parâmetros forenses calculados para os marcadores InDel (30 <i>loci</i> ) e STR (16 <i>loci</i> ) estudados na população imigrante brasileira a residir na região de Lisboa.					
HLD88	0,373	0,627	0,167	0,37	0,95

HLD101	0,387	0,613	0,191	0,37	1,01
HLD67	0,372	0,628	0,148	0,37	0,91
HLD83	0,349	0,651	0,136	0,37	0,88
HLD114	0,384	0,616	0,198	0,37	1,03
HLD48	0,418	0,582	0,195	0,35	1,02
HLD124	0,360	0,640	0,157	0,37	0,93
HLD122	0,401	0,599	0,218	0,37	1,07
HLD125	0,369	0,631	0,154	0,37	0,92
HLD64	0,406	0,594	<b><u>0,103</u></b>	0,34	<b><u>0,81</u></b>
HLD81	0,406	0,594	<b><u>0,239</u></b>	0,37	<b><u>1,12</u></b>
HLD136	0,401	0,599	0,160	0,35	0,94
HLD133	0,366	0,634	0,167	0,37	0,95
HLD97	<b><u>0,338</u></b>	<b><u>0,662</u></b>	0,115	<b><u>0,38</u></b>	0,84
HLD40	0,368	0,632	0,170	0,37	0,96
HLD128	0,375	0,625	0,188	0,37	1,00
HLD39	0,394	0,606	0,202	0,37	1,03
HLD84	0,383	0,617	0,151	0,36	0,92
<hr/>					
				<b><i>Marcadores STR</i></b>	
D3S1358	0,086	0,914	0,573	0,75	2,34
D19S433	0,108	0,892	0,608	0,72	2,56
D2S1338	<b><u>0,026</u></b>	<b><u>0,974</u></b>	0,740	<b><u>0,88</u></b>	3,92
D22S1045	<b><u>0,110</u></b>	<b><u>0,890</u></b>	0,508	<b><u>0,70</u></b>	1,99
D16S539	0,071	0,929	0,573	0,77	2,34
D18S51	0,028	0,972	0,755	0,87	4,17
D1S1656	0,055	0,945	0,652	0,81	2,91
D10S1248	0,089	0,911	0,551	0,73	2,21
D2S441	0,104	0,896	0,541	0,72	2,16
TH01	0,101	0,899	0,459	0,72	1,78
vWa	0,069	0,931	0,529	0,77	2,09
D21S11	0,080	0,920	<b><u>0,454</u></b>	0,77	<b><u>1,76</u></b>
D12S391	0,031	0,969	0,699	0,87	3,39
D8S1179	0,059	0,941	0,707	0,80	3,47
FGA	0,033	0,967	0,718	0,86	3,62
SE33	0,038	0,962	<b><u>0,770</u></b>	0,87	<b><u>4,44</u></b>

Relativamente aos marcadores InDel, os valores da MP encontram-se entre 0,338 (HLD97) e 0,430 (HLD70), já para os *loci* STR, esses valores situam-se entre 0,026 (D2S1338) e 0,110 (D22S1045). Quanto ao PD e PE dos marcadores InDel, os valores máximos e mínimos encontram-se, respetivamente, entre 0,570 (HLD70) e 0,662 (HLD97) para o PD, e entre 0,103 (HLD64) e 0,239 (HLD81) para o PE. Todos os *loci* STR apresentam valores de PD superiores a 0,880, variando entre 0,890 (D22S1045) e 0,974 (D2S1338). No que se refere ao PE, os valores situam-se entre 0,454 (D21S11) e 0,770 (SE33). Os valores de PIC para os *loci* InDel, situam-se entre 0,33 (HLD70) e 0,38 (HLD97), e entre 0,88 (D2S1338) e 0,70 (D22S1045) para os STR. A diferença nos valores de PIC entre os conjuntos de marcadores genéticos, é um claro indicador da desigualdade no grau de polimorfismo dos *loci* InDel – bialélicos – e STR. Assim, apesar dos valores constatarem uma baixa diversidade alélica, estão de acordo com o esperado para marcadores bialélicos, cujo valor de PIC não será superior a 0,38. Finalmente, em relação aos valores de TPI, os marcadores bialélicos apresentam valores entre 0,81 (HLD64) e 1,12 (HLD81), enquanto para os STR esses valores situam-se entre 1,76 (D21S11) e 4,44 (SE33). Este parâmetro quando igual ou superior a 1, é indicativo de parentesco entre os indivíduos em estudo, pelo que *loci* que apresentam esses valores, são bons marcadores para o estudo do parâmetro em questão. Os resultados obtidos neste estudo, são semelhantes aos de outros autores [22,91,100,101]

Resumidamente, os *loci* InDel, à exceção do *locus* HLD70, que exibe os valores mínimos para três dos cinco parâmetros forenses calculados, como marcadores bialélicos, apresentam resultados satisfatórios para serem utilizados em identificação genética humana, como método complementar aos STRs.

## Capítulo 4 – Conclusão

O presente trabalho teve como objetivo principal, a caracterização genética da população imigrante do Brasil residente em Lisboa, recorrendo a marcadores genéticos do tipo InDel e STRs do sistema ESS. Neste estudo, procurou-se ainda perceber, a partir das distâncias genéticas, as semelhanças e diferenças entre esta população, a população portuguesa e populações imigrantes PALOPs – Angola, Guiné-Bissau, Moçambique, São Tomé e Príncipe e Cabo Verde –, também muito representativas em termos numéricos.

Em relação aos marcadores InDel, à exceção do *locus* HLD97, todos os *loci* encontram-se em HWE, após correção de Bonferroni. Para alguns *loci*, verificou-se a ocorrência de algum desbalanceamento nas frequências alélicas, pois, tratando-se de marcadores bialélicos, deveriam estar próximas de 0,500. Este desequilíbrio nas frequências alélicas, em que um alelo se apresenta mais frequente do que o outro, ambos pertencentes ao mesmo *locus*, poderá permitir a sua utilização como marcador informativo de ancestralidade (AIM), associando um indivíduo a uma região geográfica, salvaguardando que os *loci* estudados não foram concebidos para esse fim, mas sim para identificação genética individual. No presente estudo, os *loci* HLD70 e HLD64, são os que apresentam maiores diferenças entre frequências alélicas. Contudo, a utilização de *loci* como marcadores informativos de ancestralidade, requer mais análise estatística, nomeadamente a análise de variação intrapopulacional, e estudo de um maior número de indivíduos, inclusive de diferentes regiões geográficas.

A deteção das microvariantes nos *loci* HLD97, HLD56 e HLD99, na população estudada, permite aumentar o poder de discriminação do painel de marcadores do *kit Investigator® DIPplex* reforçando, contudo, tal como sugerido por Fondevila e colaboradores, a necessidade de reformulação do *primer reverse* do *locus* HLD97, para uma correta obtenção dos perfis genéticos. Até ao momento, e para nosso melhor conhecimento, nenhum outro estudo reportou a existência de uma microvariante para o *locus* HLD56, pelo que se deveria proceder à sua sequenciação, para melhor caracterizá-la.

O painel de marcadores InDel, como *loci* bialélicos, à exceção do *locus* HLD70, apresentaram resultados satisfatórios para serem utilizados em identificação genética humana, como método complementar aos STRs.

No que se refere aos marcadores STR, todos apresentaram elevado polimorfismo, com valores de  $H_o$  entre 0,749 (D22S1045) e 0,951 (SE33). Adicionalmente, todos os *loci* encontravam-se em HWE, com exceção do *locus* TH01. No que toca aos parâmetros forenses, todos os valores de PD foram superiores a 0,880 e os valores de PIC superiores a 0,70, com exceção do *locus* D22S1045. Conclui-se assim, que o sistema *multiplex* que inclui todos os *loci* que compõem o ESS, inclusive os últimos cinco mini-STRs - D12S391 e D1S1656, D2S441, D10S1248 e D22S1045 -, é altamente informativo e representa uma importante ferramenta em identificação humana forense, bem como em testes de paternidade.

Os resultados genéticos obtidos, permitiram-nos confirmar a existência de diferenças genéticas entre a população Portuguesa e populações imigrantes oriundas do Brasil, Angola, Cabo Verde, São Tomé e Príncipe, Moçambique e Guiné-Bissau, que se apresentam assim como uma fonte de variabilidade

genética. Mais ainda, os resultados alcançados, refletem o impacto da história, ao longo dos séculos, de cada uma dessas regiões, na sua constituição genética.

Após correção de Bonferroni, os resultados revelaram a presença de LD entre seis pares de *loci*. Em cada um dos casos, os *loci* encontram-se em cromossomos diferentes, pelo que a associação alélica estará mais relacionada com o tamanho da amostra ou com a existência de subestrutura populacional, característica comum de populações miscigenadas, tal como a população Brasileira [65].

A caracterização genética da população imigrante do Brasil residente em Lisboa, pela determinação das frequências alélicas dos marcadores do tipo InDel e STR, irá permitir calcular com maior rigor a frequência do genótipo de um indivíduo da referida população, que compareça ao INMLCF- Sul, no âmbito de uma investigação forense.

## Referências Bibliográficas

- [1] J.K. Pickrell, D. Reich, Toward a new history and geography of human genes informed by ancient

- DNA, Trends Genet. 30 (2014) 377–389. doi:10.1016/j.tig.2014.07.007.
- [2] J.H. Relethford, Genetic evidence and the modern human origins debate, (2008) 555–563. doi:10.1038/hdy.2008.14.
  - [3] L. Campos, S. Canavezes, Introdução à Globalização, (2007) 166. <http://hdl.handle.net/10174/2468>.
  - [4] A.J. Spector, Globalization Or Imperialism? Neoliberal Globalization In The Age Of Capitalist Imperialism, Int. Rev. Mod. Sociol. 33 (2007) 7–26. doi:10.1007/BF00590987.
  - [5] M. Czaika, H. De Haas, The globalization of migration: Has the world become more migratory?, Int. Migr. Rev. 48 (2014) 283–323. doi:10.1111/imre.12095.
  - [6] K. Basu, Globalization , Poverty , and Inequality : What is the Relationship ? What Can Be Done ?, 34 (2006) 1361–1373. doi:10.1016/j.worlddev.2005.10.009.
  - [7] European Commission., Statistical Office of the European Communities., Push and pull factors of international migration : a comparative report, 2000. doi:10.1017/CBO9781107415324.004.
  - [8] P. Stankov, Deregulation, economic growth and growth acceleration, 2010. <http://mpira.ub.uni-muenchen.de/26485/1/Wp424.pdf>.
  - [9] U. Nations, International Migration Report 2013, United Nations Dep. Econ. Soc. Aff. Popul. Div. (2013) 1. doi:ST/ESA/SER.A/346.
  - [10] United Nations, International Migration Report 2015, (2016). [http://www.un.org/en/development/desa/population/migration/publications/migrationreport/docs/MigrationReport2015\\_Highlights.pdf](http://www.un.org/en/development/desa/population/migration/publications/migrationreport/docs/MigrationReport2015_Highlights.pdf).
  - [11] D. Studies, Crossing the Atlantic Frontier Brazilian Migrants in Portugal, (2015).
  - [12] C. Rego, M.F. Mendes, J.M.R. Santos, M.G. Magalhães, Perfil dos imigrantes em Portugal: dos países de origem às regiões de destino, Actas Do 15º Congr. Da Assoc. Port. Desenvolv. Reginal 6 a 11 Julho. (2009) 17–39.
  - [13] P. Seguro, 7172 TÍTULO II Dos portugueses no Brasil e dos brasileiros em Portugal, (2000) 7172–7187.
  - [14] B. Digitais, U.C. Digitalis, U.C. Pombalina, U.C. Impactum, B. Digitais, U.C. Digitalis, O novo estatuto de igualdade entre Brasileiros e Portugueses Autor ( es ): Publicado por : URL persistente : DOI : Accessed : Ribeiro , Jeancezar Ditz de Souza Universidade Católica de Petrópolis URI : <http://hdl.handle.net/10316.2/33941> DOI : <http://dx.doi.org/10.1016/j.tig.2014.07.007>, (2018).
  - [15] SEF, Relatório de Imigração Froteiras e Asilo, (2016).
  - [16] Sistema de Segurança Interna [SSI], Relatório Anual de Segurança Interna 30 de março de 2015, Relatório Anu. Segurança Interna – Ano. (2015) 254. [http://www.ansr.pt/InstrumentosDeGestao/Documents/Relatório Anual de Segurança Interna \(RASI\)/RASI 2016.pdf](http://www.ansr.pt/InstrumentosDeGestao/Documents/Relatório Anual de Segurança Interna (RASI)/RASI 2016.pdf).
  - [17] P. Dias, R. Machado, A.P. Ferreira, G. de E.P. e F. SEF, Relatório de Imigração, Fronteiras e Asilo 2017, (2017) 170. [http://sefstat.sef.pt/Docs/Rifa\\_2014.pdf](http://sefstat.sef.pt/Docs/Rifa_2014.pdf).
  - [18] A. Esteves, Os Efeitos Da Crise Económica Na Situação Laboral Dos Imigrantes : O Caso Dos Brasileiros Em Portugal, (2015) 135–152.

- [19] B. Fausto, A escravidão - índios e negros, História Do Bras. (1996) 28–31. [http://www.edusp.com.br/detlivro.asp?ID=35179%5Cnhttp://educaleaks.dominiotemporario.com/doc/Historia\\_do\\_Brasil\\_Boris\\_Fausto\\_Resumo.pdf%5Cnhttp://www.caccto.com.br/material/d00044/Material\\_6\\_E-MED\\_2A\\_195641.pdf%5Cnhttp://www.historiaemperspectiva.com/2012](http://www.edusp.com.br/detlivro.asp?ID=35179%5Cnhttp://educaleaks.dominiotemporario.com/doc/Historia_do_Brasil_Boris_Fausto_Resumo.pdf%5Cnhttp://www.caccto.com.br/material/d00044/Material_6_E-MED_2A_195641.pdf%5Cnhttp://www.historiaemperspectiva.com/2012).
- [20] Apex - Brasil, Investment Guide to Brazil, 2017. doi:10.1017/CBO9781107415324.004.
- [21] S.S. Goes Filho, As Fronteiras do Brasil, 2013.
- [22] V.R. Da Costa Aguiar, E. De Vargas Wolfgramm, F.S.V. Malta, A.G. Bosque, A.D.C. Mafia, V.C. De Oliveira Almeida, F. De Andrade Caxito, V.C. Pardini, A.C.S. Ferreira, I.D. Louro, Updated Brazilian STR allele frequency data using over 100,000 individuals: An analysis of CSF1PO, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, FGA, Penta D, Penta E, TH01, TPOX and vWA loci, Forensic Sci. Int. Genet. 6 (2012) 504–509. doi:10.1016/j.fsigen.2011.07.005.
- [23] Q. Cordeiro, B.R. Souza, H. Correa, C. Guindalini, M.H. Hutz, H. Vallada, M.A. Romano-silva, A review of psychiatric genetics research in the Brazilian population, Rev Bras Psiquiatr. 31 (2009) 154–162. doi:10.1590/S1516-44462009000200013.
- [24] K.L.P. WILL, Genocídio Indígena no Brasil, 2014.
- [25] M. Hutchison, S. Nichols, M. Santos, H. Onsrud, S. Paixao, Demarcation and registration of indigenous lands in Brazil, (2006) 46.
- [26] O. Maia, Aldeias e missões nas capitanias do Ceará e Rio Grande: catequese, violência e rivalidades 1, 19 (1980).
- [27] P. Martín, O. García, B. Heinrichs, I. Yurrebaso, A. Aguirre, A. Alonso, Population genetic data of 30 autosomal indels in Central Spain and the Basque Country populations, Forensic Sci. Int. Genet. 7 (2013) 27–30. doi:10.1016/j.fsigen.2012.10.003.
- [28] IBGE, Censo Demográfico 2010. Características da População e dos Domicílios., Inst. Bras. Geogr. E Estatística. (2010) 48. doi:0104-3145.
- [29] V. Miranda, A resurgence of black identity in Brazil? Evidence from an analysis of recent censuses, Demogr. Res. 32 (2015) 1603–1630. doi:10.4054/DemRes.2015.32.59.
- [30] Q. Cordeiro, B.R. Souza, H. Correa, C. Guindalini, M.H. Hutz, H. Vallada, M.A. Romano-Silva, A review of psychiatric genetics research in the Brazilian population TT - Uma revisão sobre a pesquisa genética psiquiátrica na população brasileira, Rev Bras Psiquiatr. 31 (2009) 154–162. [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1516-44462009000200013](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1516-44462009000200013).
- [31] R. Chakraborty, Human Population Genetics: Drift and Migration, Life Sci. (2001) 1–5. doi:10.1002/9780470015902.a0001786.pub2.
- [32] B.J.B. Keats, S.L. Sherman, Population Genetics, Sixth Edit, Elsevier, 2013. doi:10.1016/B978-0-12-383834-6.00015-X.
- [33] J.F. Crow, Population genetics history: a personal view., Annu. Rev. Genet. 21 (1987) 1–22. doi:10.1146/annurev.genet.21.1.1.
- [34] J.M. Butler, Fundamentals of Forensic DNA Typing, 2009.
- [35] D.J. Schaid, S.J. Jacobsen, Biased Tests of Association: Comparisons of Allele Frequencies when



- Departing from Hardy-Weinberg Proportions, *Am. J. Epidemiol.* 149 (1999) 706–711. <http://aje.oxfordjournals.org/content/149/8/706.abstract>.
- [36] B.J.B. Keats, S.L. Sherman, Population Genetics, in: Emery Rimoin's *Princ. Pract. Med. Genet.* (Sixth Ed., 6th ed., 2013: pp. 1–12. doi:10.1111/j.1558-5646.2007.00164.x.
  - [37] A. Hastings, Hardy – Weinberg Theorem, *Life Sci.* (2001) 1–4.
  - [38] O. Mayo, A century of Hardy-Weinberg equilibrium, *Twin Res. Hum. Genet.* 11 (2008) 249–256. doi:10.1375/twin.11.3.249.
  - [39] B. Chen, J.W. Cole, C. Grond-Ginsbach, Departure from Hardy Weinberg Equilibrium and genotyping error, *Front. Genet.* 8 (2017) 1–6. doi:10.3389/fgene.2017.00167.
  - [40] J.M. Butler, *DNA Databases*, Elsevier Inc., 2012. doi:10.1016/B978-0-12-374513-2.00008-7.
  - [41] W. Bateson, R. Saunders, Er; Punnett, Experimental studies in the physiology of heredity, *Reports to Evol. Comm. R. Soc.* 2 (1905) 1–55, 80–99. doi:10.1007/BF01975751.
  - [42] T. H. Morgan, Sex Limited Inheritance in *Drosophila*, *Science* (80-. ). XXXII (1910) 120–122.
  - [43] T. H. Morgan, Chromosomes and Heredity, *Am. Nat.* XLIV (1910) 449–496.
  - [44] K. Lobo, Ingrid; Shaw, *Discovery and Types of Genetic Linkage*, (2008). <https://www.nature.com/scitable/topicpage/discovery-and-types-of-genetic-linkage-500> (accessed March 7, 2018).
  - [45] D.& J.E. Hartl, *Genetics: Principles and Analysis*, 1998.
  - [46] M. Dawn Teare, J.H. Barrett, Genetic linkage studies, *Lancet.* 366 (2005) 1036–1044. doi:10.1016/S0140-6736(05)67382-5.
  - [47] M. Slatkin, Linkage disequilibrium — understanding the evolutionary past and mapping the medical future, *Nat. Rev. Genet.* 9 (2008) 477–485. doi:10.1038/nrg2361.Linkage.
  - [48] J. Van Der Werf, Basics of Linkage and Gene Mapping, *QTL Course Belo Horiz. - Bras.* (2000) 45–54.
  - [49] A. Amorim, *Genética Forense*, *Acad. Das Ciências Lisboa.* 1 (2015) 21.
  - [50] M.A. Jobling, P. Gill, Encoded evidence: DNA in forensic analysis, *Nat. Rev. Genet.* 5 (2004) 739–751. doi:10.1038/nrg1455.
  - [51] E. Murphy, *Forensic DNA Typing*, (2017) 42–45.
  - [52] R. Enfsi, D.N. a Working, G. April, *Dna-Database Management Review and Recommendations*, (2013) 1–85.
  - [53] L. Roewer, DNA fingerprinting in forensics: Past, present, future, *Investig. Genet.* 4 (2013) 1. doi:10.1186/2041-2223-4-22.
  - [54] M.L. Phillips, Crime Scene Genetics: Transforming Forensic Science through Molecular Technologies, *Bioscience.* 58 (2008) 484. doi:10.1641/B580604.
  - [55] M. Vennemann, A. Koppelkamm, mRNA profiling in forensic genetics I: Possibilities and limitations, *Forensic Sci. Int.* 203 (2010) 71–75. doi:10.1016/j.forsciint.2010.07.006.



- [56] E.K. Hanson, J. Ballantyne, Circulating MicroRNAs, 1024 (2013) 221–234. doi:10.1007/978-1-62703-453-1.
- [57] J.L. Weber, D. David, J. Heil, Y. Fan, C. Zhao, G. Marth, Human Diallelic Insertion / Deletion Polymorphisms, (2002) 854–862.
- [58] R.E. Mills, C.T. Luttig, C.E. Larkins, A. Beauchamp, C. Tsui, W.S. Pittard, S.E. Devine, An initial map of insertion and deletion (INDEL) variation in the human genome, *Genome Res.* 16 (2006) 1182–1190. doi:10.1101/gr.4565806.
- [59] J.M. Butler, *Advanced Topics in Forensic DNA Typing: Methodology*, 2012. doi:10.1111/j.1556-4029.2012.02199.x.
- [60] C.V. Da Silva, S. Matos, H. Afonso Costa, P. Morais, R.M. Dos Santos, R. Espinheira, J. Costa Santos, A. Amorim, Genetic portrait of south Portugal population with InDel markers, *Forensic Sci. Int. Genet.* 7 (2013) 101–103. doi:10.1016/j.fsigen.2013.03.009.
- [61] R. Pereira, C. Phillips, C. Alves, A. Amorim, Á. Carracedo, L. Gusmão, Insertion/deletion polymorphisms: A multiplex assay and forensic applications, *Forensic Sci. Int. Genet. Suppl. Ser.* 2 (2009) 513–515. doi:10.1016/j.fsigss.2009.09.005.
- [62] F. Reis, C. Vieira da Silva, A. Amorim, V. Bogas, T. Ribeiro, M.J. Porto, H. Afonso Costa, Genetic characterization of the Brazilian immigrant population in Lisboa with InDel genetic markers (P), *Forensic Sci. Int. Genet. Suppl. Ser.* (2017) 0–1. doi:10.1016/j.fsigss.2017.09.123.
- [63] V. Samuel, *Forensic Identification by Using Insertion- deletion Polymorphisms*, (2015).
- [64] Qiagen, *Investigator DIPplex Handbook Sample & Assay Technologies* QIAGEN Sample and Assay Technologies, (2014).
- [65] T.D.J.B. Ferreira Palha, E.M. Ribeiro Rodrigues, G.C. Cavalcante, A. Marrero, I.R. De Souza, C.J. Seki Uehara, C.H.A. Silveira Da Motta, D. Koshikene, D.A. Da Silva, E.F. De Carvalho, G. Chemale, J.M. Freitas, L. Alexandre, R.T.F. Paranaíba, M.P. Soler, S. Santos, Population genetic analysis of insertion-deletion polymorphisms in a Brazilian population using the Investigator DIPplex kit, *Forensic Sci. Int. Genet.* 19 (2015) 10–14. doi:10.1016/j.fsigen.2015.03.015.
- [66] J.M.J.M. Butler, *Forensic DNA typing: biology, technology, and genetics of STR markers*, 2005. doi:10.1007/s13398-014-0173-7.2.
- [67] D. Decanine, O papel de marcadores moleculares na genética forense, *Bras. Crimin.* 5 (2016) 18–27. doi:10.15260/rbc.v5i2.123.
- [68] V.C. Tucker, A.J. Hopwood, C.J. Sprecher, R.S. McLaren, D.R. Rabbach, M.G. Ensenberger, J.M. Thompson, D.R. Storts, Developmental validation of the PowerPlex® ESI 16 and PowerPlex® ESI 17 Systems: STR multiplexes for the new European standard, *Forensic Sci. Int. Genet.* 5 (2011) 436–448. doi:10.1016/j.fsigen.2010.09.004.
- [69] P. Corporation, *PowerPlex(R) ESI 17 System*, n.d.
- [70] P.M. Schneider, The ESS loci Expansion of the European Standard Set of DNA Database Loci — the Current Situation the ESS loci, *Profiles DNA.* (2009) 6–7.
- [71] P. Gill, R. Sparkes, G. Tully, *DNA Profiling in Forensic Science*, *Encycl. Life Sci.* (2001) 1–8. www.els.net.
- [72] P.S. Walsh, D.A. Metzger, R. Higuchi, *Biotechniques 30th anniversary gem Chelex 100 as a*

medium for simple extraction of DNA for PCR-based typing from forensic material, *Biotechniques*. 54 (2013) 506–513. doi:10.2144/000114018.

- [73] ThermoFisher, Scientific, Quantifiler™ HP and Trio DNA Quantification Kits, (2017).
- [74] J.W. Jorgenson, K.D. Lukacs, Free-zone electrophoresis in glass capillaries., *Clin. Chem.* 27 (1981) 1551–1553.
- [75] ThermoScientific, DNA Fragment Analysis by Capillary Electrophoresis, Thermo Fish. Handbooks. (2012) 1–220. doi:4474504.
- [76] L. Excoffier, H.E.L. Lischer, Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows, *Mol. Ecol. Resour.* 10 (2010) 564–567. doi:10.1111/j.1755-0998.2010.02847.x.
- [77] S. Kumar, G. Stecher, K. Tamura, MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets., *Mol. Biol. Evol.* 33 (2016) msw054. doi:10.1093/molbev/msw054.
- [78] M. Fondevila, C. Phillips, C. Santos, R. Pereira, L. Gusmão, A. Carracedo, J.M. Butler, M. V. Lareu, P.M. Vallone, Forensic performance of two insertion-deletion marker assays, *Int. J. Legal Med.* 126 (2011) 725–737. doi:10.1007/s00414-012-0721-7.
- [79] B.L. LaRue, J. Ge, J.L. King, B. Budowle, A validation study of the Qiagen Investigator DIPplex kit; An INDEL-based assay for human identification, *Int. J. Legal Med.* 126 (2012) 533–540. doi:10.1007/s00414-012-0667-9.
- [80] S.L. Friis, C. Børsting, E. Rockenbauer, L. Poulsen, S.F. Fredslund, C. Tomas, N. Morling, Forensic Science International : Genetics Typing of 30 insertion / deletions in Danes using the first commercial indel, 6 (2012) 72–74. doi:10.1016/j.fsigen.2011.08.002.
- [81] A. Zidkova, A. Horinek, V. Kebrdlova, M. Korabecna, Application of the new insertion-deletion polymorphism kit for forensic identification and parentage testing on the Czech population, *Int. J. Legal Med.* 127 (2013) 7–10. doi:10.1007/s00414-011-0649-3.
- [82] K.M. Seong, J.H. Park, Y.S. Hyun, P.W. Kang, D.H. Choi, M.S. Han, K.W. Park, K.W. Chung, Population genetics of insertion-deletion polymorphisms in South Koreans using Investigator DIPplex kit, *Forensic Sci. Int. Genet.* 8 (2014) 80–83. doi:10.1016/j.fsigen.2013.06.013.
- [83] L. Poulsen, M.S. Farzad, C. Børsting, C. Tomas, V. Pereira, N. Morling, Population and forensic data for three sets of forensic genetic markers in four ethnic groups from Iran: Persians, Lurs, Kurds and Azeris, *Forensic Sci. Int. Genet.* 17 (2015) 43–46. doi:10.1016/j.fsigen.2015.03.010.
- [84] G. Hefke, S. Davison, M.E. D’Amato, Forensic performance of Investigator DIPplex indels genotyping kit in native, immigrant, and admixed populations in South Africa, *Electrophoresis*. 36 (2015) 3018–3025. doi:10.1002/elps.201500243.
- [85] C.R.B. Hill, J.M. Butler, The Highly Polymorphic STR Locus SE33 : Concordance , & Population Variation, (2011).
- [86] R. Rodenbusch, S. Schumacher, A.Z. Gastaldo, F.G.C. Lanes, L.P. Maciel, T. Grandi, Candice Tosi Michelin, Cíntia Costi, C.M.D. da Silva, Population genetic data for 11 STR loci, including SE33, in Southern Brazil, *Leg. Med.* (2009).
- [87] Promega, PowerPlex® Fusion 6C System for Use on the Genetic Analyzers, 2017.
- [88] A.A. Westen, T. Kraaijenbrink, E.A. Robles De Medina, J. Harteveld, P. Willemse, S.B. Zuniga, K.J.

- Van Der Gaag, N.E.C. Weiler, J. Warnaar, M. Kayser, T. Sijen, P. De Knijff, Comparing six commercial autosomal STR kits in a large Dutch population sample, *Forensic Sci. Int. Genet.* 10 (2014) 55–63. doi:10.1016/j.fsigen.2014.01.008.
- [89] V.R.C. Aguiar, A.M. de Castro, V.C.O. Almeida, F.S.V. Malta, A.C.S. Ferreira, I. D. Louro, New CODIS core loci allele frequencies for 96,400 Brazilian individuals, (2014). doi:10.1016/j.fsigen.2014.04.015.
- [90] M.F. Álvarez, Forensic Performance of Short Amplicon Insertion-Deletion ( InDel ) Markers, 22nd Internatinal Symp. Hum. Identif. (2011).
- [91] S.R.R. Torres, C.J.S. Uehara, A.F. Sutter-Latorre, B.S. De Almeida, T.S. Sauerbier, Y.C.N. Muniz, A.R. Marrero, I.R. De Souza, Population genetic data and forensic parameters of 30 autosomal InDel markers in Santa Catarina State population, Southern Brazil, *Mol. Biol. Rep.* 41 (2014) 5429–5433. doi:10.1007/s11033-014-3414-0.
- [92] W. Du, Z. Peng, C. Feng, B. Zhu, B. Wang, Forensic efficiency and genetic variation of 30 InDels in Vietnamese and Nigerian populations, 8 (2017) 88934–88940.
- [93] J. Vansina, New Linguistic Evidence and the Bantu Expansion Author ( s ): J . Vansina Reviewed work ( s ): Published by : Cambridge University Press, 36 (2012) 173–195.
- [94] Joana Gorjão Henriques, SÃO TOMÉ E PRÍNCIPE : A escravatura durou até à independência, *Rev. Público.* (2016). <https://acervo.publico.pt/mundo/noticia/em-sao-tome-e-principe-a-escravatura-durou-ate-a-independencia-1729886> (accessed July 6, 2018).
- [95] Joana Gorjão Henriques, CABO VERDE : Um hipermercado de escravos, *Rev. Público.* (2016). <https://acervo.publico.pt/mundo/noticia/de-cabo-verde-a-angola-na-rota-da-escravatura-em-cinco-paises-1729858> (accessed July 6, 2018).
- [96] Joana Gorjão Henriques, GUINÉ-BISSAU : Alimentar o comércio de escravos de Cabo Verde, *Rev. Público.* (2016). <https://acervo.publico.pt/mundo/noticia/-guinebissau-alimentava-o-comercio-de-escravos-de-cabo-verde-1729883> (accessed July 6, 2018).
- [97] Joana Gorjão Henriques, ANGOLA : O grande “produtor” de escravos, *Rev. Público.* (2016). <https://acervo.publico.pt/mundo/noticia/-angola-o-grande-produtor-de-escravos-1729882> (accessed July 6, 2018).
- [98] Joana Gorjão Henriques, MOÇAMBIQUE : A escravatura não desaparece com tratados, *Rev. Público.* (2016). [https://www.bbc.com/portuguese/reporterbbc/story/2007/05/070406\\_escravosmocambiqu e2ss.shtml](https://www.bbc.com/portuguese/reporterbbc/story/2007/05/070406_escravosmocambiqu e2ss.shtml) (accessed July 6, 2018).
- [99] C. Glycério, S. Silvia, Moçambique deixou marca genética, mas pouca herança, *BBC.* (2007). [https://www.bbc.com/portuguese/reporterbbc/story/2007/05/070406\\_escravosmocambiqu e2ss.shtml](https://www.bbc.com/portuguese/reporterbbc/story/2007/05/070406_escravosmocambiqu e2ss.shtml) (accessed July 7, 2018).
- [100] T. De Jesus, B. Ferreira, E. Martins, R. Rodrigues, G. Chaves, A. Marrero, I. Rainha, D. Souza, C. Julien, S. Uehara, C. Henrique, A. Silveira, D. Koshikene, D. Aparecida, E. Fagundes, D. Carvalho, G. Chemale, J.M. Freitas, L. Alexandre, R.T.F. Paranaíba, M. Perruccio, S. Santos, Forensic Science International : Genetics Population genetic analysis of insertion – deletion polymorphisms in a Brazilian population using the Investigator DIPplex kit, *Forensic Sci. Int. Genet.* 19 (2015) 10–14. doi:10.1016/j.fsigen.2015.03.015.
- [101] A. Amorim, R. Marques dos Santos, C. Vieira da Silva, H. Afonso Costa, R. Espinheira, J. Costa

Santos, Genetic portrait of Brazilian immigrant population living in Lisboa, *Forensic Sci. Int. Genet.* 6 (2012) 121–124. doi:10.1016/j.fsigen.2012.01.008.

## Capítulo 5 – Anexos

### 5.1 Características dos *loci* incluídos nos Sistemas *Investigator® DIPplex* e *PowerPlex® ESI 17*

Tabela 5.1 Pannel de 30 marcadores, 29 *loci* InDel e a Amelogenina, incluídos no *kit Investigator® DIPplex (QIAGEN)* e pannel de 17 marcadores, 16 *loci* STR e Amelogenina, presentes no *kit PowerPlex® ESI 17 (Promega)*. Para o primeiro *kit*, constam da tabela o nome dos marcadores, localização no cromossoma, sequência de inserção/deleção - HLD do inglês, *Human Locus DIP* – e respetiva referência na base de dados SNPforID. Adaptado de [64]. Para o segundo *kit*, constam o nome dos marcadores, localização no cromossoma, categoria e motivo de repetição e o N° - número - de alelos. Adaptado de [69].

<i>kit Investigator® DIPplex</i>				<i>kit PowerPlex® ESI 17</i>			
Locus DIP	Localização no cromossoma	Sequência	ID SNPforID	Locus	Localização no cromossoma	Categoria e motivo de repetição	N° alelos
Amelogenina X	Xp22.1-22.3	-	M55418	D22S1045	22q12.3	[ATT]n	7-20
Amelogenina Y	Yp11.2	-	M55419	D2S1338	2q35	[TGCC]n[TTCC]n	10-28
HLD77	7q31.1	TAAG	rs1611048	D19S433	19q12	Complexa – [AAGG]n	5.2-18.2
HLD45	2q31.1	CACG	rs2307959	D3S1358	3p21.31	Complexa – [TCTA]n	9-20
HLD131	7q36.2	TGGGCTTATT	rs1611001	D2S441	2p14	[TCTA]n	8-17
HLD70	6q16.1	AGCA	rs2307652	D10S1248	10q26.3	[GGAA]n	8-19
HLD6	16q13	GCAGGACTGGGCACC	rs1610905	D1S1656	1q42	Complexa – [TAGA]n	9-20.3
HLD111	17p11.2	CACA	rs1305047	D18S51	18q21.33	[AGAA]n	7-27
HLD58	5q14.1	AGGA	rs1610937	D16S539	16q24.1	[GATA]n	4-16
HLD56	4q25	TAAAGT	rs2308292	D12S391	12q	Complexa – [AGAT]n[AGAC]n	14-27
HLD118	20p11.1	CCCCA	rs16438	D21S11	21q21.1	Complexa – [TCTA]n	24-38
HLD92	11q22.2	GTTT	rs17174476	vWA	12p13.31	Complexa – [TCTA]n	10-24
HLD93	12q22	ACTTT	rs2307570	TH01	11p15.5	[AATG]n	3-13.3
HLD99	14q23.1	TGAT	rs2308163	SE33	6q14	Complexa – [AAAG]n	4.2-42
HLD88	9q22.32	CCACAAAGA	rs8190570	FGA	4q28	Complexa – [TTTC]n	14-50.2
HLD101	15q26.1	GTAG	rs2307433	D8S1179	8q24.13	Complexa – [TCTA]n	7-9
HLD67	5q33.2	CTACTGAC	rs1305056	Amelogenina	Xp22.1-22.3 e Y	-	X, Y
HLD83	8p22	AAGG	rs2308072				
HLD114	17p13.3	TCCTATTCTACTCTGAAT	rs2307581				
HLD48	2q11.2	GACTT	rs28369942				
HLD124	22q12.3	GTGGA	rs6481				
HLD122	21q22.11	GAAGTCTGAGG	rs8178524				
HLD125	22q11.23	ATTGCC	rs16388				
HLD64	5q12.3	GACAAA	rs1610935				
HLD81	7q21.3	GTAAGCATTGT	rs17879936				
HLD136	22q13.1	TGTTT	rs16363				
HLD133	3p22.1	CAACCTGGATT	rs2067235				
HLD97	13q12.3	AGAGAAAGCTGAAG	rs17238892				
HLD40	1p32.3	GGACAGTGGCCACTAGGAG	rs2307956				
HLD128	1q31.3	ATTAAATA	rs2307924				
HLD39	1p22.1	CCTAAACAAAAATGGGAT	rs17878444				
HLD84	8q24.12	CTTTC	rs3081400				

## 5.2 Resultados obtidos na quantificação do DNA das amostras

Tabela 5.2 Resultados obtidos na quantificação do DNA das amostras estudadas. Da tabela constam os códigos das amostras e respetivo resultado da quantificação do DNA em ng/μL. As amostras que, aquando da quantificação, não foi detetado DNA, apresentam-se com hífen.

Código Amostra	(ng/μL)	Código Amostra	(ng/μL)	Código Amostra	(ng/μL)	Código Amostra	(ng/μL)	Código Amostra	(ng/μL)
GENPOP1001	0,28	GENPOP1051	0,18	GENPOP1104	0,15	GENPOP1159	0,13	GENPOP1217	0,26
GENPOP1002	0,46	GENPOP1053	0,18	GENPOP1105	0,16	GENPOP1160	0,42	GENPOP1218	0,48
GENPOP1003	0,17	GENPOP1054	0,82	GENPOP1106	0,27	GENPOP1161	0,66	GENPOP1219	0,20
GENPOP1004	0,05	GENPOP1055	0,42	GENPOP1107	0,05	GENPOP1162	0,10	GENPOP1220	0,84
GENPOP1005	0,31	GENPOP1056	0,45	GENPOP1108	0,02	GENPOP1163	0,04	GENPOP1221	0,65
GENPOP1006	0,38	GENPOP1057	0,63	GENPOP1110	0,04	GENPOP1164	0,03	GENPOP1222	0,69
GENPOP1007	0,18	GENPOP1058	0,28	GENPOP1111	0,04	GENPOP1165	0,03	GENPOP1223	0,30
GENPOP1008	0,23	GENPOP1059	0,36	GENPOP1112	0,02	GENPOP1166	0,06	GENPOP1224	0,50
GENPOP1009	0,34	GENPOP1060	0,32	GENPOP1113	0,07	GENPOP1167	0,05	GENPOP1225	0,63
GENPOP1010	0,14	GENPOP1061	0,22	GENPOP1114	0,29	GENPOP1168	0,36	GENPOP1226	0,39
GENPOP1011	0,56	GENPOP1062	0,44	GENPOP1115	0,11	GENPOP1169	0,27	GENPOP1227	0,92
GENPOP1012	0,46	GENPOP1063	0,49	GENPOP1116	0,06	GENPOP1170	0,10	GENPOP1228	0,65
GENPOP1013	0,41	GENPOP1064	0,25	GENPOP1117	0,15	GENPOP1171	0,10	GENPOP1229	0,31
GENPOP1014	0,33	GENPOP1065	0,40	GENPOP1118	0,08	GENPOP1172	0,20	GENPOP1230	0,41
GENPOP1015	0,45	GENPOP1066	0,47	GENPOP1119	-	GENPOP1173	0,08	GENPOP1231	0,60
GENPOP1016	0,35	GENPOP1067	0,16	GENPOP1120	0,08	GENPOP1174	0,20	GENPOP1232	0,42
GENPOP1017	0,61	GENPOP1068	0,35	GENPOP1121	0,04	GENPOP1175	0,34	GENPOP1233	0,49
GENPOP1018	0,75	GENPOP1069	0,52	GENPOP1122	0,12	GENPOP1176	0,59	GENPOP1234	0,85
GENPOP1019	0,25	GENPOP1070	0,56	GENPOP1123	0,02	GENPOP1177	0,02	GENPOP1235	0,62
GENPOP1020	0,39	GENPOP1071	0,23	GENPOP1124	0,06	GENPOP1178	2,82	GENPOP1236	0,27
GENPOP1021	0,39	GENPOP1072	0,24	GENPOP1125	0,04	GENPOP1179	0,44	GENPOP1237	0,30
GENPOP1022	0,26	GENPOP1073	0,44	GENPOP1126	0,02	GENPOP1180	0,19	GENPOP1238	0,80
GENPOP1023	0,12	GENPOP1074	0,31	GENPOP1127	0,02	GENPOP1181	2,15	GENPOP1239	1,09
GENPOP1024	0,35	GENPOP1075	0,32	GENPOP1128	0,10	GENPOP1182	0,40	GENPOP1240	0,24
GENPOP1025	0,54	GENPOP1076	0,39	GENPOP1129	0,07	GENPOP1183	0,22	GENPOP1241	0,79
GENPOP1026	0,68	GENPOP1077	0,45	GENPOP1130	0,16	GENPOP1184	0,09	GENPOP1242	0,48
GENPOP1027	0,22	GENPOP1078	0,49	GENPOP1131	0,14	GENPOP1185	0,16	GENPOP1243	0,92
GENPOP1028	0,25	GENPOP1079	0,50	GENPOP1132	0,17	GENPOP1186	0,42	GENPOP1244	0,25
GENPOP1029	0,34	GENPOP1080	0,16	GENPOP1133	0,07	GENPOP1188	0,52	GENPOP1245	0,23
GENPOP1030	0,43	GENPOP1081	0,35	GENPOP1134	0,03	GENPOP1189	0,54	GENPOP1246	0,26
GENPOP1031	0,45	GENPOP1082	0,39	GENPOP1135	0,19	GENPOP1190	0,25	GENPOP1247	0,75
GENPOP1032	0,67	GENPOP1083	0,60	GENPOP1136	0,05	GENPOP1192	0,29	GENPOP1248	0,29
GENPOP1033	0,30	GENPOP1084	0,40	GENPOP1137	0,18	GENPOP1198	0,13	GENPOP1249	1,05
GENPOP1034	0,50	GENPOP1087	0,44	GENPOP1138	0,15	GENPOP1199	0,14	GENPOP1250	0,36
GENPOP1035	0,35	GENPOP1088	0,29	GENPOP1139	0,14	GENPOP1200	0,56	GENPOP1251	0,46
GENPOP1036	0,44	GENPOP1089	0,27	GENPOP1141	0,20	GENPOP1201	0,20	GENPOP1252	0,61
GENPOP1037	0,26	GENPOP1090	0,34	GENPOP1142	0,23	GENPOP1202	0,89	GENPOP1253	0,52
GENPOP1038	0,21	GENPOP1091	0,14	GENPOP1143	0,13	GENPOP1203	0,17	GENPOP1254	0,28
GENPOP1039	0,71	GENPOP1092	0,57	GENPOP1144	0,11	GENPOP1204	0,51	GENPOP1255	0,98
GENPOP1040	0,36	GENPOP1093	-	GENPOP1145	0,10	GENPOP1205	0,34	GENPOP1256	1,19
GENPOP1041	0,89	GENPOP1094	0,01	GENPOP1146	0,05	GENPOP1206	0,74	GENPOP1257	0,58
GENPOP1042	0,41	GENPOP1095	0,04	GENPOP1147	0,22	GENPOP1207	0,26	GENPOP1258	0,48
GENPOP1043	0,26	GENPOP1096	0,08	GENPOP1148	0,15	GENPOP1208	0,72	GENPOP1259	0,54
GENPOP1044	0,35	GENPOP1097	0,01	GENPOP1151	0,14	GENPOP1209	0,33	GENPOP1260	0,84
GENPOP1045	0,39	GENPOP1098	0,12	GENPOP1152	0,14	GENPOP1210	0,44		
GENPOP1046	0,44	GENPOP1099	0,09	GENPOP1153	0,10	GENPOP1211	0,34		
GENPOP1047	0,47	GENPOP1100	0,10	GENPOP1154	0,06	GENPOP1213	0,37		
GENPOP1048	0,42	GENPOP1101	0,19	GENPOP1156	0,04	GENPOP1214	0,24		
GENPOP1049	0,87	GENPOP1102	0,04	GENPOP1157	0,05	GENPOP1215	0,43		
GENPOP1050	0,68	GENPOP1103	0,18	GENPOP1158	0,07	GENPOP1216	0,21		

### 5.3 Representação gráfica das frequências alélicas para 16 loci do kit PowerPlex® ESI 17

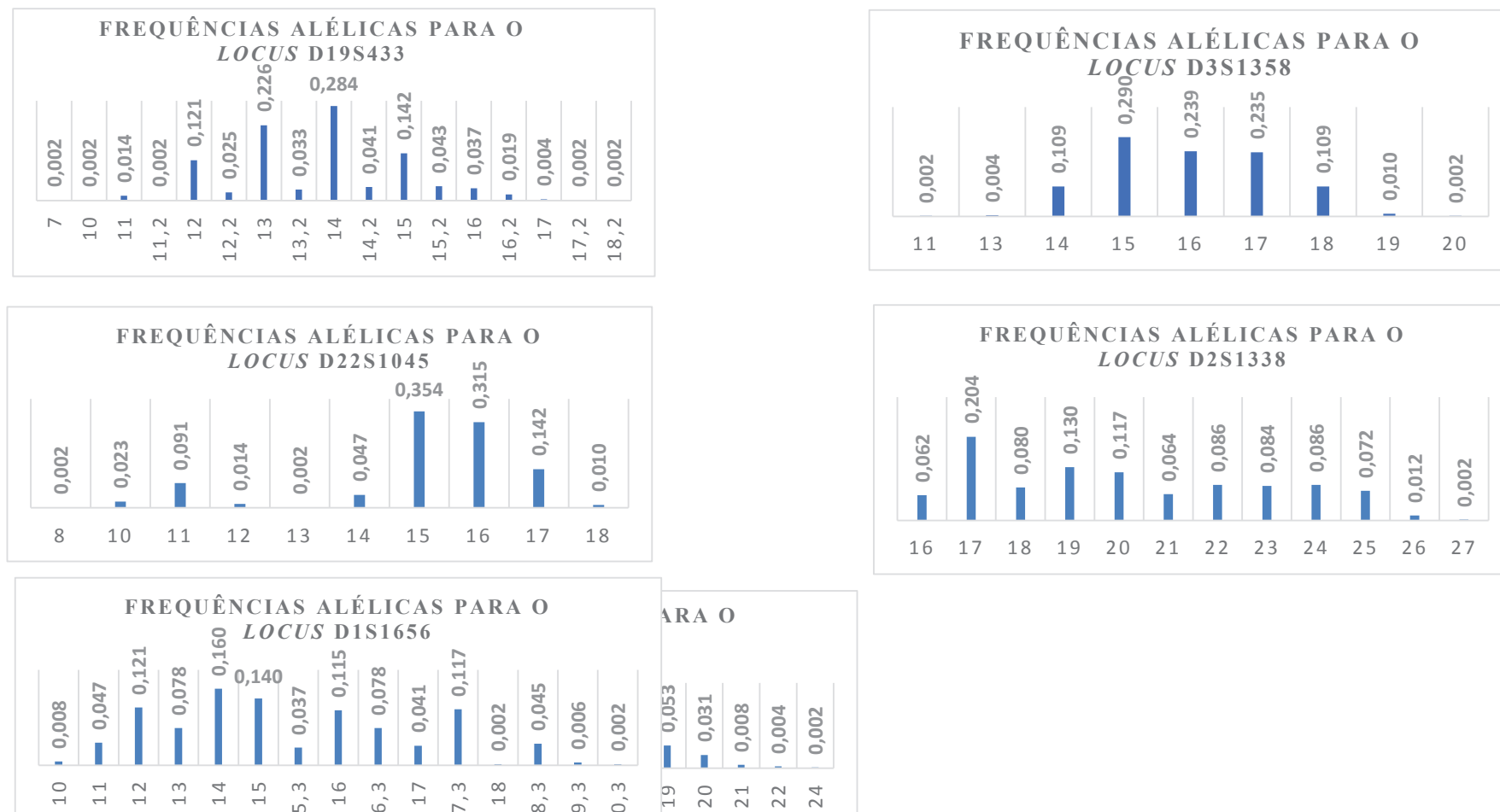


Figura 5.1 Representação gráfica das frequências alélicas para 16 loci do kit PowerPlex® ESI 17 (Promega) na população Brasileira residente em Lisboa.



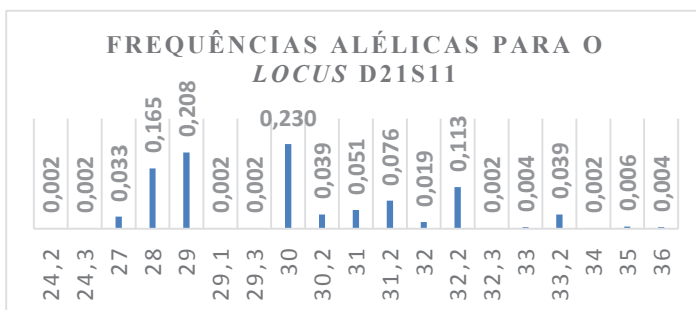
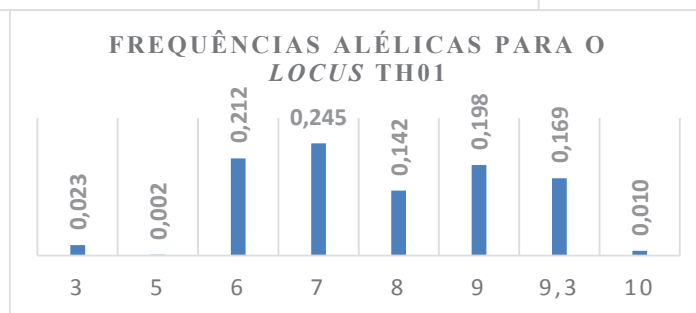
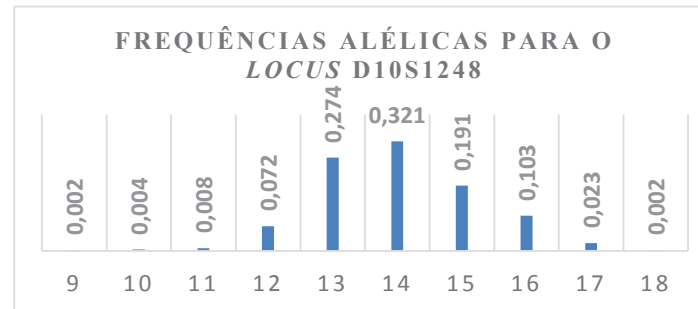
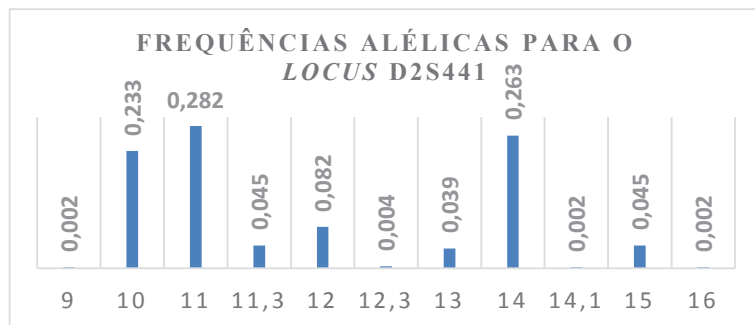
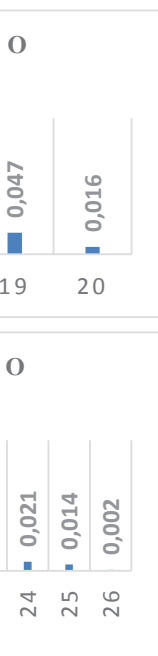


Figura 5.1 – Continuação – Representação gráfica das frequências alélicas para 16 loci do kit PowerPlex® ESI 17 (Promega) na população Brasileira residente em Lisboa.

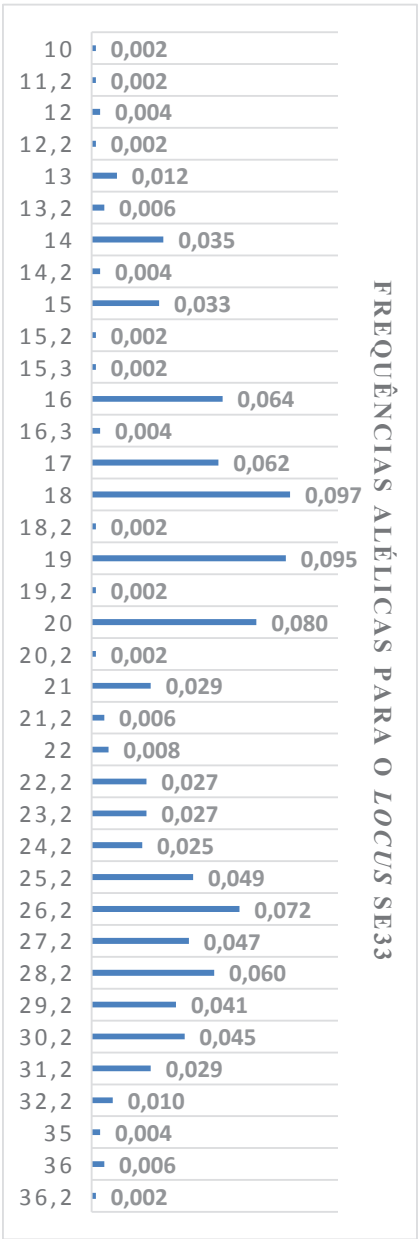
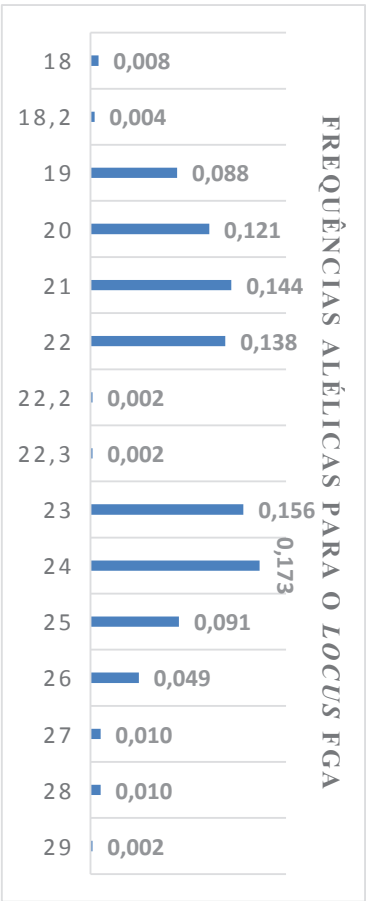
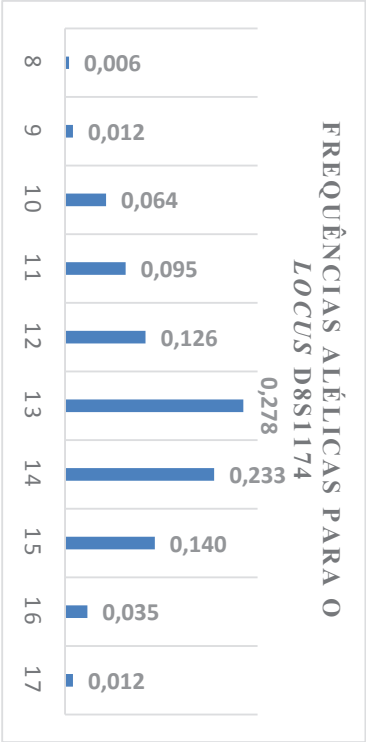


Figura 5.1 – Continuação – Representação gráfica das frequências alélicas para 16 loci do kit PowerPlex® ESI 17 (Promega) na população Brasileira residente em Lisboa

#### 5.4 Resultados da análise de LD entre os 46 marcadores em estudo

Tabela 5.3 **Resultados obtidos aquando da análise de LD para os 46 loci em estudo:** pares de *loci* assinalados a cinzento, após correção Bonferroni, não se apresentam em LD ( $0,05 > p < 0,000048$ ), a cinzento escuro estão os pares de *loci* que, mesmo após a correção, mantêm-se em LD ( $p < 0,000048$ ).

[illegible]